

UNIVERSITE MONTPELLIER II  
SCIENCES ET TECHNIQUES DU LANGUEDOC

# THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE MONTPELLIER II

*Discipline : Biologie des populations*

**AUX ORIGINES DE LA DIVERSITE DE LA PATATE DOUCE (*IPOMOEA BATATAS*)**

*Une enquête phylogéographique en Amérique tropicale (aire d'origine)  
et en Océanie (aire d'introduction)*

Présentée par

**Roullier Caroline**

*Ecole doctorale : Systèmes Intégrés en Biologie, Agronomie, Géosciences, Hydrosciences et  
Environnement (SIBAGHE)*

JURY

**M. Gepts Paul** (PR, Université de Californie, Davis)  
**Mme Heyer Evelyne** (PR, Muséum National d'Histoire Naturelle de Paris)  
**Mme Elias Marianne** (CR, Muséum National d'Histoire Naturelle de Paris)  
**Mme Di Piazza Anne** (CR, CNRS-CREDO, Marseille)  
**M. David Jacques** (PR, INRA, Montpellier)  
**M. Vincent Lebot** (CR, CIRAD, Montpellier)  
**M. McKey Doyle** (PR, Université Montpellier 2)

Rapporteur  
Rapporteuse  
Examinatrice  
Examinatrice  
Examineur  
Directeur de thèse  
Co-directeur de thèse

# Remerciements

Tout d'abord, je tiens à remercier vivement les bailleurs de fonds qui ont contribué au financement de ce projet de recherche : en premier lieu le « Fonds Pacifique », géré par le Secrétariat Permanent pour le Pacifique et inscrit au budget du Ministère des Affaires étrangères (MAE), mais aussi le projet SYNTHESYS (Systematic Resources, [www.synthesys.info](http://www.synthesys.info)) financé par la Communauté Européenne, le CIRAD, le CNRS et enfin l'Ecole Normale Supérieure de Cachan.

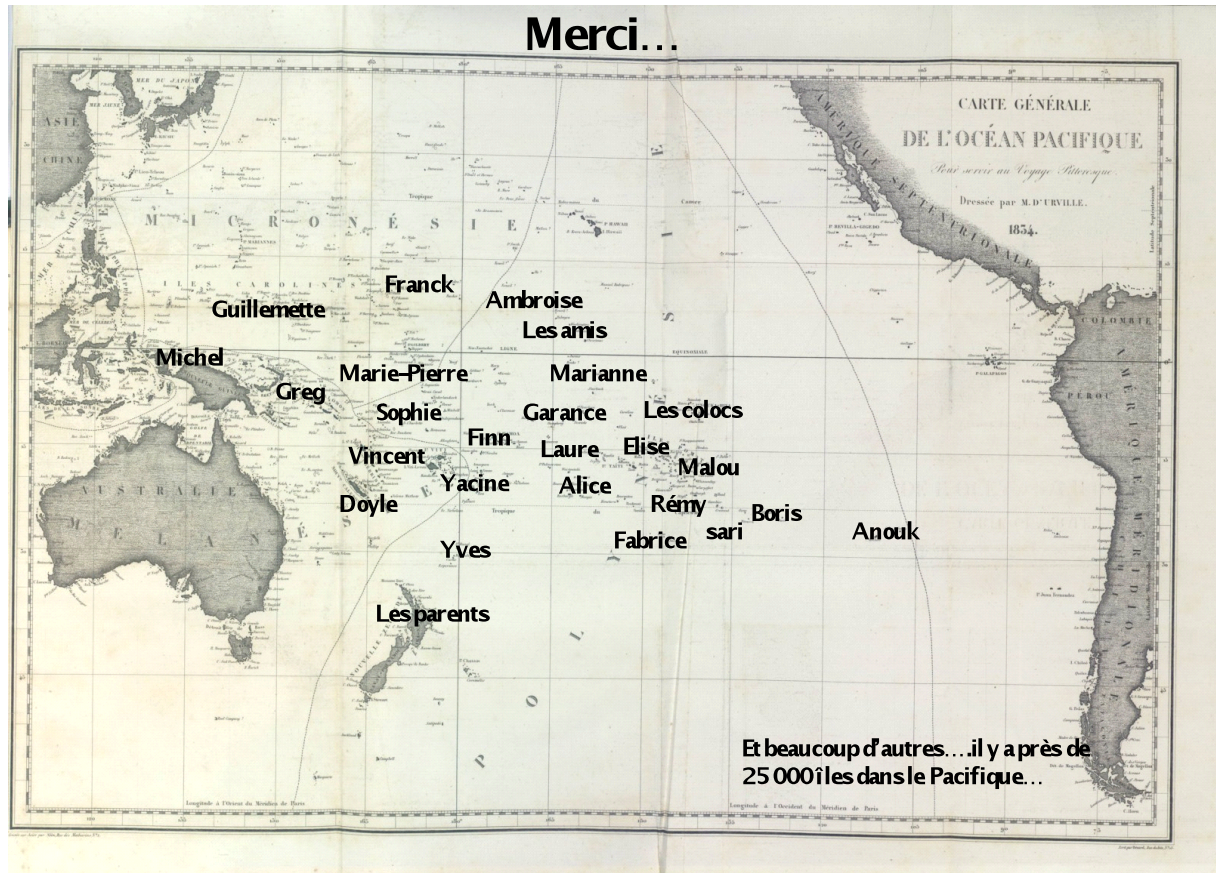
Je souhaite aussi remercier chaleureusement tous les organismes qui ont fourni le matériel végétal nécessaire à cette vaste étude génétique. Il y a tout d'abord les collections *ex situ* : le Centre International de la Pomme de Terre (CIP, Lima, Pérou), l'Institut National de Recherche Agricole de Papouasie Nouvelle Guinée (NARI, Papouasie Nouvelle Guinée), le Centre Technique de la Recherche Agricole du Vanouatou (CTRAV, Vanouatou), le Service de Développement Rural (SDR, Papeete, Tahiti), l'Institut National des Sciences Agronomiques (NIAS, Tsukuba, Japon), l'Institut National des Cultures (NICS, Tsukuba, Japon), et aussi les différents herbiers nationaux, l'herbier du Muséum d'Histoire Naturelle de Londres (BM, Royaume-Uni), l'herbier des Jardins Royaux de Kew (K, Royaume-Uni), l'herbier du Bishop Museum (BISH, Hawaii, Etats-Unis), et l'herbier national de l'Université de Leiden (L, Pays-Bas).





Une thèse, en un sens, peut ressembler à une longue traversée du Pacifique. Je tiens à remercier tous ceux qui ont initié puis accompagné cette aventure. Telle une myriades d'îles, ils m'ont permis de faire escale et de trouver la force et le désir de poursuivre le voyage.

**Merci...**



# Table des matières

<b>TABLE DES MATIERES</b>	<b>1</b>
<b>TABLE DES FIGURES ET TABLEAUX</b>	<b>6</b>
<b>CHAPITRE 1</b>	<b>9</b>
<b>INTRODUCTION GENERALE.....</b>	<b>9</b>
<b>I. DOMESTIQUER DES PLANTES .....</b>	<b>10</b>
<b>I.1. Spatialité et temporalité de la domestication .....</b>	<b>10</b>
<i>I.1.1.Des centres aux non-centres</i>	<i>10</i>
<i>I.1.2. Domestication unique ou multi-locale ?</i>	<i>11</i>
<i>I.1.3. Temporalité</i>	<i>12</i>
<b>I.2. Le(s) processus de domestication.....</b>	<b>13</b>
<i>I.2.1. Syndromes de domestication</i>	<i>13</i>
<i>I.2.2. Forces à l'œuvre</i>	<i>14</i>
<i>I.2.3. Domestiquer des plantes annuelles propagées par graines : l'exemple des céréales</i>	<i>15</i>
Bases génétiques de la domestication.	15
D'un modèle de transition rapide à une transition prolongée	16
<i>I.2.4. Domestiquer et reproduire une plante par propagation clonale: l'exemple des plantes à racines et tubercules tropicales</i>	<i>16</i>
Pourquoi avoir choisi de propager certaines plantes par voie végétative?	17
L'exemple du manioc dans les systèmes agricoles amazoniens	18
L'igname africaine	19
<b>I.3. Polypléidie et domestication .....</b>	<b>19</b>
<i>I.3.1. Formation des polypléides</i>	<i>20</i>
<i>I.3.2. Avantages évolutifs de la polypléidie dans le contexte de l'évolution des plantes cultivées</i>	<i>21</i>
<b>II. DIFFUSER DES PLANTES : LA DIVERSIFICATION DES PLANTES DOMESTIQUEES HORS DE LEUR FOYER D'ORIGINE.....</b>	<b>22</b>
<b>II.1. La diffusion : un aperçu.....</b>	<b>22</b>
<i>II.1.1. Modes de diffusion</i>	<i>22</i>

II.1.2. <i>Echelles spatio-temporelles</i>	23
II.1.3. <i>Conséquences génétiques et évolutives de la diffusion des plantes domestiquées</i>	23
<b>II.2. La diffusion des hommes et des plantes en Océanie.....</b>	<b>25</b>
II.2.1. <i>Une origine des plantes océaniques plutôt sur Sabul que Sunda</i>	25
II.2.2. <i>Les premiers migrants du Pléistocène en "Océanie proche"</i>	25
II.2.3. <i>De la cueillette aux débuts de l'agriculture en Nouvelle-Guinée</i>	25
II.2.4. <i>Des réseaux d'interactions précoces avec les îles d'Asie du Sud-est</i>	26
II.2.5. <i>L'arrivée des austronésiens en Océanie proche et le peuplement de l'Océanie lointaine</i>	27
L'arrivée des austronésiens	27
La colonisation en deux étapes de l'"Océanie lointaine"	28
II.2.6. <i>Des contacts préhistoriques entre les Polynésiens et les Américains?</i>	28
Le poulet asiatique en Amérique	29
Des plantes cultivées américaines en Océanie	29
Les routes empruntées	30
II.2.7. <i>Des plantes sur des pirogues</i>	31
Un gradient de diversité d'Ouest en Est	31
Des introductions multiples	32
Re-domestiquer localement certaines espèces	32
Diffusion, propagation clonale et stérilité	32
II.2.8. <i>Pratiques paysannes et diversification des plantes à multiplication végétative dans les systèmes vivriers océaniques : rôles de la mutation somatique et de la reproduction sexuée</i>	33
II.2.9. <i>L'arrivée puis l'installation des Européens</i>	35
Introduction et re-distribution des plantes cultivées	35
Les systèmes vivriers fragilisés	36
 <b>CHAPITRE 2</b>	 <b>38</b>
<b>Modèle biologique, problématiques &amp; outils de l'enquête .....</b>	<b>38</b>
 <b>I. ORIGINES ET DOMESTICATION DE LA PATATE DOUCE : FAITS ET HYPOTHESES .....</b>	 <b>39</b>
<b>I.1. Position taxonomique .....</b>	<b>39</b>
I.1.1. <i>Le genre Ipomoea</i>	39
I.1.2. <i>Le complexe ou série Batatas</i>	40
<b>I.2. Aux origines de <i>I. batatas</i> : retour sur une controverse .....</b>	<b>40</b>
<b>I.3. L'origine botanique de <i>I. batatas</i> : deux scénarios en jeu.....</b>	<b>41</b>
<b>I.4. Origine géographique de la domestication .....</b>	<b>41</b>
<b>I.5. Le caractère domestiqué : la racine tubéreuse.....</b>	<b>42</b>

I.6. Biologie de la reproduction et sources de variabilité .....	42
II. DIFFUSION ET DIVERSIFICATION DE LA PATATE DOUCE EN OCEANIE.....	43
II.1. L'histoire des introductions de la patate douce en Océanie : L'hypothèse tripartite, un résumé .....	43
II.1.1. <i>Les introductions préhistoriques ou la route Kumara</i> .....	44
II.1.2. <i>Les introductions historiques : les routes Camote et Batata</i> .....	45
II.2. Diversités océaniques ; une histoire d'adoption, d'adaptation, et de diversification .....	46
II.3. A la recherche de preuves génétiques.....	47
III. PROBLEMATIQUE(S) ET PLAN DE THESE .....	48
IV. LES OUTILS DE L'ENQUETE .....	50
IV.1. Les plantes .....	50
IV.2. Les marqueurs .....	51
IV.3. Analyses génétiques des polyploïdes à partir de marqueurs microsatellites nucléaires .....	52
CHAPITRE 3 .....	54
Article : Caractérisation de la structure de la diversité génétique de la patate douce dans son aire d'origine .....	54
Combining chloroplast and nuclear microsatellites to investigate origin and dispersal of New World sweet potato landraces .....	54
CHAPITRE 4 .....	68
Article : Relations phylogénétiques entre <i>Ipomoea batatas</i> et les espèces sauvages apparentées .....	68
Disentangling the origins of cultivated sweet potato ( <i>Ipomoea</i> <i>batatas</i> (L.) Lam.) .....	68
CHAPITRE 5 .....	95
Article : Origine de la diversité de la patate douce ( <i>Ipomoea</i> <i>batatas</i> (L.) Lam.) en Nouvelle Guinée, un second centre de diversité .....	96
On the origin of sweet potato ( <i>Ipomoea batatas</i> (L.) Lam.) genetic diversity in New Guinea, a secondary centre of diversity .....	96

<b>CHAPITRE 6</b>	<b>121</b>
Article : Les collections historiques révèlent des patrons de diffusion de la patate douce ( <i>Ipomoea batatas</i> (L.) Lam.) en Océanie brouillé par les mouvements et la recombinaison des plantes modernes.....	121
Historical collections reveal patterns of diffusion of sweet potato in Oceania obscured by modern plant movements and recombination.....	121
<b>CHAPITRE 7</b>	<b>138</b>
Discussion générale.....	138
<b>I. AUTO OU ALLOPOLYPLOIDIE ? DE LA DIFFICULTE DE DEFINIR LA NATURE DES POLYPLOÏDES .....</b>	<b>139</b>
<b>II. QUELLES RELATIONS ENTRE I. TRIFIDA ET I. BATATAS ? DE LA DIFFICULTE D'IDENTIFIER ET DE DELIMITER DES ESPECES .....</b>	<b>141</b>
<b>III. DOMESTICATION UNIQUE OU MULTI-LOCALE ? .....</b>	<b>142</b>
<b>IV. A LA RECHERCHE DE PREUVES GENETIQUES DE L'HYPOTHESE TRIPARTITE .....</b>	<b>143</b>
IV.1. Les variétés polynésiennes et la voie Kumara .....	143
IV.2. D' hier à aujourd' hui .....	144
IV.3. A l' Ouest du Pacifique.....	144
IV.4. D' Est en Ouest, d' Ouest en Est.....	145
<b>V. CREER DE LA DIVERSITE CHEZ UNE PLANTE A PROPAGATION CLONALE.....</b>	<b>146</b>
V.1. Fonctionnement(s) des systèmes mixtes.....	146
V.2. Les facteurs qui influencent le fonctionnement des systèmes mixtes .....	147
V.2.1. La biologie et l'écologie de la reproduction de la plante	147
V.2.2. Les caractéristiques « génomiques »	147
V.2.3. Les conditions de l'environnement	148
V.2.4. Les pratiques agricoles et la relation culturelle à la « diversité »	148

<b>VI. RACONTER L'HISTOIRE DES HOMMES EN OCEANIE AU TRAVERS DE CELLES DES ESPECES COMMENSALES OU DOMESTIQUES : DES PISTES DE REFLEXION.....</b>	<b>149</b>
VI.1. La nécessité d'une approche temporelle.....	150
VI.2. Le difficile choix des marqueurs moléculaires.....	151
VI.3. La nécessité d'une approche transversale .....	151
<b>VII. QUELLES IMPLICATIONS POUR LA GESTION DES RESSOURCES GENETIQUES? .....</b>	<b>151</b>
VII.1. La conservation et l' amélioration des plantes à racines et tubercules.....	151
VII.2. Quelle diversité « utile » pour cette espèce ?.....	152
VII.3. De la possibilité de mettre en œuvre des méthodes de gestion à la ferme et d' amélioration participative pour la patate douce en Océanie .....	154
VII.4. Ressources génétiques et appropriation du vivant.....	155
<b>RÉFÉRENCES</b>	<b>156</b>

# Table des figures et tableaux

- Figure 1-1** : Principaux foyers de domestication des plantes
- Figure 1-2** : Phénotypes comparés de formes sauvage et cultivée
- Figure 1-3** : Cycle de vie du manioc domestiqué dans les systèmes de culture amérindiens.
- Figure 1-4** : Voies de formation d'un autopolyploïde et mode de ségrégation polysomique associée
- Figure 1-5** : Histoire évolutive du coton.
- Figure 1-6** : Les principales plantes du complexe horticole océanien.
- Figure 1-7** : Les terres émergées Sunda et Sahul (en vert) il y a environ 50,000 ans lorsque les premiers migrants du pléistocène ont peuplé la Nouvelle Guinée depuis l'Asie.
- Figure 1-8** : Vagues de peuplement de l'Océanie et mouvements post-colonisation des Polynésiens.
- Figure 1-9** : L'histoire croisée des Hommes et des plantes en Océanie.
- Figure 2-1** : Aperçu de la production mondiale de patate douce.
- Figure 2-2** : Espèces d'*Ipomoea* mentionnées dans cette étude
- Figure 2-3** : Possibilités d'hybridation entre les différentes espèces du complexe *Batatas*.
- Figure 2-4** : Les deux scénarios proposés pour expliquer l'origine du génome hexaploïde de la patate douce
- Figure 2-5** : Eléments de reproduction sexuée de la patate douce.
- Figure 2-6** : Les différentes routes d'introduction de la patate douce en Océanie telles que définies dans l'hypothèse tripartite
- Figure 2-7** : Spécimens de patates douces observés lors du premier voyage de J. Cook dans le Pacifique en 1769.
- Figure 2-8** : Jardins de patate douce.
- Figure 3-1** : Geographic distribution of CIP sweet potato accessions used in the present study.
- Figure 3-2** : Distribution of chloroplast haplotypes in New World sweet potato landraces.
- Figure 3-3** : Nuclear genetic diversity distribution of American sweet potato landraces.
- Figure 3-4** : Factorial analysis of the 130 accessions genotyped with nuclear SSRs.
- Figure 4-1** : Geographical distribution of the sampling.
- Figure 4-2** : Genetic relationships of *I. batatas*, five wild relatives and *Ipomoea* sp. accessions based on chloroplast DNA analyses.
- Figure 4-3** : Genetic relationships of *I. batatas*, five wild relatives and *Ipomoea* sp. accessions based on nuclear DNA analyses.
- Figure 4-4** : Taxa boundaries of *I. batatas*, *I. trifida* and polyploid *Ipomoea* sp. accessions assessed using two complementary clustering approaches (SSRs dataset).
- Figure 4-5** : Geographical patterns of cpDNA lineages and nuclear clusters (STRUCTURE results) of *I. batatas*, *I. trifida* and polyploid *Ipomoea* sp. accessions.
- Figure 5-1** : Geographic provenience of the 417 New Guinean sweet potato landraces from the NARI and NIAS collections, and definition of the regional and agro-ecosystem groups used in our study.
- Figure 5-2** : Patterns of genetic differentiation of sweet potato landraces in New Guinea.
- Figure 5-3** : Geographical origin of New Guinea and Tropical America sweet potato landraces (as distinguished by chloroplast data, and nuclear data with DAPC analysis at  $K = 2$ ) and possible paths of sweet potato introduction into New Guinea.
- Figure 5-4** : Tropical American origin of the genetic diversity of New Guinean landraces as assessed by nuclear markers.
- Figure 5-5** : Allelic richness that can be captured for a given number of samples.
- Figure 6-1** : The pre-historic and historical dispersal of sweet potato in Oceania as postulated by the tripartite hypothesis.
- Figure 6-2** : Geographical distribution of nuclear and chloroplast genetic variation in sweet potato modern sampling.
- Figure 6-3** : Global patterns of genetic differentiation.

**Figure 6-4 :** Genetic constitution of herbarium specimens collected from the 18th century to the early 20th

**Figure 6-5 :** Hierarchical representation of the radial tree

**Figure 6-6 :** Frequency distribution of genetic dissimilarity based on the calculation of all pairwise Manhattan distances between genotypes within landraces of Western Pacific and Polynesia.

**Tableau 1-1 :** Liste des principales plantes du complexe horticole Lapitas

**Tableau 2-1 :** Les 14 espèces d'*Ipomoea* de la série *Batatas*, leur niveau de ploïdie et leur aire de répartition.

**Tableau 2-2 :** Logiciels d'analyses de génétique des populations permettant de traiter des données microsatellites polyploïdes

**Table 3-1 :** Characterization of the seven cpSSR markers used in our study and the polymorphism observed for them.

**Table 3-2 :** Characterization of the nuclear SSRs used in our study.

**Table 3-3 :** Multilocus haplotypes defined with seven chloroplast markers

**Table 3-4 :** Geographical distribution of cpSSR haplotypes accessed by country and resumed in two geographical regions, the Northern and the Southern regions.

**Table 3-5 :** Nuclear diversity (number of alleles) of the Southern and Northern regions counting respectively 49 and 81 accessions.

**Table 4-1 :** Contingency table confronting cpDNA haplotype lineages with Structure clusters among *I. batatas*, *I. trifida*, *Ipomoea* sp. and *I. tabascana*.

**Table 4-2 :** Genetic diversity of the four geographically well-sampled taxa as revealed by nuclear SSRs.

**Table 5-1 :** Nuclear and chloroplast diversity within regional and agroecosystem-type groups in New Guinea.

**Table 5-2 :** Analysis of molecular variance (AMOVA) of nuclear SSRs within New Guinea (among agroecosystem groups, among regional groups and within regional groups) and among tropical America (Southern and Northern genepools) and New Guinea.

**Table 5-3 :** Comparison of nuclear and chloroplast diversity between tropical America and New Guinea.

**Table 5-4 :** Distance and differentiation between accessions from New Guinea (highland and lowland groups) and the neotropics (Northern and Southern genepools), measured by intra- and inter-group pairwise distances (Lynch distances) above the diagonal and  $F_{st}$  values below the diagonal.

**Table 6-1 :** Comparison of nuclear and chloroplast diversity between tropical America and Oceania.

**Table 6-2 :** Multilocus clonal lineages shared between regions for the modern sample

**Encadré 1:** Espaces et environnements Océaniens



*Toute la vie maritime! Tout dans la vie maritime!  
Dans mon sang s'insinue cette séduction subtile  
Et je rêve de voyages sur un mode indéterminé  
Ah! Les lignes des côtes lointaines, aplaties par l'horizon!  
Ah! Les caps, les îles, les plages sablonneuses!  
Les solitudes maritimes, comme certains moments dans le Pacifique  
Où par l'effet de je ne sais quelle suggestion apprise à l'école  
On sent peser sur les nerfs le fait que c'est là le plus grand des Océans,  
Et le monde, et la saveur des choses deviennent un désert au dedans de nous!  
L'étendue plus humaine, plus éclaboussée, de l'Atlantique!  
L'Indien, de tous les océans le plus mystérieux!  
O Méditerranée, douce, sans aucun mystère, classique,  
ô mer  
Bonne à lécher des esplanades regardaient par des statues blanches  
en des jardins proches!*

*Toutes les mers, tous les détroits, toutes les baies, tous les golfes  
je voudrais les serrer sur ma poitrine, bien les sentir, et mourir!*

Fernando Pessoa  
Extrait d'Ode maritime

*Aux origines de la patate douce*



# *CHAPITRE 1*

## *INTRODUCTION GENERALE*

*Domestiquer et diffuser des plantes :  
Comment s'est façonnée la diversité des plantes cultivées ?*



# I. DOMESTIQUER DES PLANTES

Dans l'histoire des sociétés humaines, une étape fondamentale fut très probablement celle du passage de l'appropriation de ressources spontanées à celle de la production de ressources domestiquées (Barrau 1978). En d'autres termes, le passage d'une économie de prédation à celle d'une économie de production (Childe 1942). Depuis ses origines à la fin du Paléolithique, l'homme moderne, *Homo sapiens*, pratiquait des modes de prédation divers, cueillette, pêche ou chasse, et se déplaçait au gré de ses besoins et des variations de son environnement. Dès 12 000 ans av. JC., de nombreux sites archéologiques attestent que certaines populations ont commencé de façon plus ou moins concomitante et indépendante à pratiquer l'agriculture et à domestiquer certaines plantes et animaux (Diamond 2002, Gepts 2004). Moins de 500 espèces de plantes auraient ainsi été domestiquées, un nombre finalement assez restreint vu la vaste diversité du monde végétal (Gepts 2012). Cette transition, souvent qualifiée de « révolution néolithique » (Childe 1942), a été porteuse de transformations majeures pour bon nombre de sociétés concernées, tant sur le plan culturel que technique, économique et politique (Smith 1976, Diamond 2002) : une explosion démographique a généralement suivi l'avènement de l'agriculture, conséquence directe de la production de nourriture en surplus, générant des vagues de colonisation et l'absorption de la plupart des groupes de chasseurs-cueilleurs environnants (Diamond et Bellwood 2003). L'avènement de l'agriculture a aussi conduit plus ou moins rapidement à la sédentarisation des populations productrices de nourriture, le développement de cités et la complexification de l'organisation sociale. Toutes les sociétés n'ont pourtant pas eu la même histoire. Les espèces à domestiquer, les conditions écologiques de leur domestication, tout comme les conditions matérielles et culturelles diffèrent grandement d'une région à l'autre (Vrydaghs et Denham 2007).

Face à cette diversité de situations, il est parfois difficile de trouver une définition consensuelle au terme de domestication (Harris 2007). La domestication d'une plante, au sens étymologique, consiste à l'intégrer au foyer, au « *domus* ». Charles Darwin dans son ouvrage « *The variation of animals and plants under domestication* » (Darwin 1868), attribue rapidement au mot « domestication » un sens biologique. La domestication est alors vue comme un processus évolutif, de sélection, qui conduit à des transformations morphologiques et génétiques de la plante que l'on essaie de cultiver (Purugganan et Fuller 2009). La culture regroupe un ensemble de tâches élémentaires, de soins apportés aux plantes dans un milieu aménagé par l'homme. L'agriculture correspond à la culture de plantes domestiquées. Ces définitions guideront notre propos ici.

## I.1. Spatialité et temporalité de la domestication

Il est couramment admis que l'agriculture est apparue indépendamment sur tous les continents et ce de façon plus ou moins concomitante dans les différentes régions du monde (Diamond 2002, Gepts 2004, Purugganan et Fuller 2009) (Figure 1-1).

### I.1.1.Des centres aux non-centres

C'est au botaniste suisse, Alphonse De Candolle (1883), qu'on doit les premiers travaux sur l'origine des plantes cultivées, dans lesquels il expose notamment l'importance de recourir à un faisceau d'arguments convergents issus de plusieurs disciplines (notamment l'archéologie, la biogéographie, la botanique, la linguistique et l'histoire) pour reconstruire l'origine des plantes cultivées. La notion de « centre d'origine » - une région circonscrite où débute l'agriculture et à partir de laquelle elle diffuse - a été relayée par l'agronome et généticien russe Nikolai Vavilov. En 1920, il organise des prospections dans la plupart des continents et rassemble d'importantes collections de variétés de plantes cultivées. Vavilov considère que la diversité des formes cultivées d'une espèce est le résultat de la sélection



humaine et reflète donc l'ancienneté de sa domestication. Suivant ce raisonnement, la région géographique où l'espèce cultivée présente le plus de diversité, à condition qu'elle contienne aussi les espèces spontanées apparentées, doit être le lieu où elle a été domestiquée initialement. Selon ces critères, il identifie huit zones ou centres d'origine, chacun partagé entre plusieurs plantes : la Chine, l'Asie du Sud/Sud-est, l'Asie centrale, le Moyen-Orient, le bassin méditerranéen, l'Éthiopie, le Mexique et l'Amérique du Sud (andine). Depuis, l'essor des données archéologiques combinées à celles obtenues par l'étude des relations génétiques entre les formes sauvages et cultivées sur la base de marqueurs neutres et sous sélection ont substantiellement modifié notre vision de l'origine géographique de la domestication (Zeder 2006, Burger et al. 2008). En premier lieu, des nouvelles régions comme la Nouvelle-Guinée (Lebot 1999, Pickersgill 2004, Denham 2003 et 2011, Allaby 2007), le Nord-est de l'Amérique (Smith 2006) ou encore le bassin amazonien (Olsen et Schall 2001, Clement et al. 2010), se sont ajoutées à la liste des centres d'origine. Parallèlement, les frontières de ces centres ont été révisées et étendues, si bien que pour bon nombre d'entre eux, il devient finalement difficile d'en délimiter un contour net (Purugganan et Fuller 2009, Price et Bar-Yosef 2011, Piperno 2011) (Figure 1-1).

Harlan (1971) propose la terminologie de « non-centres » pour caractériser de vastes régions géographiques qui auraient connu des domestications locales ici ou là, gardant le terme de centre pour des régions bien définies (Mésio-Amérique, Chine, Moyen-Orient). L'origine de la domestication des plantes dans le bassin amazonien illustre parfaitement cette idée d'une origine diffuse (Clement et al. 2010). Alors que la plupart de la diversité cultivée semble être concentrée au cœur même du bassin amazonien, là où les densités de populations d'avant la conquête étaient les plus fortes, les données phylogéographiques récentes obtenues notamment pour des espèces comme le manioc (*Manihot esculenta*) (Olsen et Schall 2001, Olsen 2004, Léotard et al. 2009), l'ananas (*Ananas comosus*) (revue par Coppens d'Eeckengrugge et Duval 2009) ou encore le pejbaye (*Bactris gasipaes*) (revue par Clement et al. 2010), montrent que la majorité des plantes d'Amazonie sont originaires de la périphérie du bassin amazonien, chacune provenant d'un endroit différent. Centres de diversité et centres d'origine ne sont donc pas nécessairement des notions équivalentes (Harlan 1992). Les plantes cultivées, lors de leur diffusion, peuvent entrer secondairement en contact avec des formes sauvages apparentées (Ellstrand et al. 1999, Weissmann et al. 2005), élargissant ainsi leurs bases génétiques initiales. Même pour les régions pour lesquelles la notion de centre reste pertinente, comme le centre mésoaméricain, il apparaît que les différentes plantes - maïs (*Zea mays* ssp. *mays*), haricots (*Phaseolus* sp.) ou encore piments (*Capsicum* sp.) - ont été domestiquées dans des zones distinctes, parfois même écologiquement différenciées (Piperno 2011).

### ***1.1.2. Domestication unique ou multi-locale ?***

Pour un certain nombre de plantes, les données génétiques disponibles mettent en évidence l'existence d'un foyer unique de domestication pour une espèce donnée, à partir duquel aurait été dispersée la forme cultivée. L'exemple le mieux caractérisé d'un tel scénario reste le maïs. Les différentes études basées sur plusieurs types de marqueurs (isozymes, gènes de domestication, microsatellites) montrent que l'ensemble des formes cultivées dérivent de populations de téosinte (*Zea mays* ssp. *parviglumis*) des basses-terres du Sud-ouest du Mexique (Doebley et al. 1984, Wang et al. 1999 et 2001, Matsuoka et al. 2002). Le manioc (Olsen et Schall 1999 et 2001), le tournesol (*Helianthus annuus*) (Wills et Burke 2006) auraient aussi une origine unique. Pour d'autres plantes, l'interprétation du signal génétique traduit plutôt l'existence de foyers multiples de domestication. Sur le continent américain, il existe plusieurs exemples de domestications multiples d'espèces proches dans différentes localités (revue par Pickersgill 2007), comme dans le genre *Cucurbita* (courges) (Sanjur et al. 2002), *Capsicum* (piments) (Pickersgill 1989) ou encore *Gossypium* (coton) (Wendel et al. 1992). De même, certaines plantes ont pu être domestiquées indépendamment à partir de populations d'une même espèce génétiquement différenciées et géographiquement isolées, comme le haricot commun (*Phaseolus vulgaris*) (Gepts 1998, Chacon et al. 2005), qui trouve son origine à la fois en Amérique centrale et en Amérique du Sud, ou encore le taro (*Colocasia esculenta*), qui a été probablement domestiqué indépendamment en Nouvelle-Guinée et en Asie du Sud-est (Lebot et Aradhya 1991, Kriek et al. 2004, Lebot et al. 2004). Il est important de souligner que l'existence de domestications multiples,

avérée par des approches de génétique ne signifie pas nécessairement que ces événements soient indépendants culturellement (Piperno 2011).

Pour certaines plantes, notamment bon nombre de céréales du Croissant Fertile ou encore le riz asiatique (*Oryza sativa*), la controverse ne tarit pas, l'un ou l'autre des scénarios étant tantôt invoqué en fonction du type de marqueurs moléculaires utilisés, de l'exhaustivité de l'échantillonnage et des analyses statistiques employées (Allaby et al. 2008, Burger et al. 2008, Glémin et al. 2009). Souvent, l'utilisation de marqueurs moléculaires couvrant largement le génome (les AFLPs ou les microsatellites) suggère une origine unique tandis que les données phénotypiques ou encore les données de séquences de gènes sous sélection au cours de la domestication sont en faveur d'une origine multiple (Allaby et al. 2008, Burger et al. 2008).

La difficulté d'inférer l'origine d'une plante cultivée réside principalement dans le fait que la similitude génétique qui existe entre formes sauvages et cultivées peut être le reflet d'une identité par descendance (transmission verticale de l'ancêtre sauvage à la forme cultivée capturée lors du processus de domestication) comme le résultat de flux de gènes secondaires (transmission horizontale) entre populations sauvages et cultivées lors de la dispersion de ces dernières (Gepts 2004). Étonnamment, les maïs cultivés des hauts plateaux mexicains sont plus proches génétiquement de la téosinte (l'ancêtre sauvage), que ne le sont les maïs cultivés des basses terres. Les maïs des hauts plateaux semblent d'ailleurs être à l'origine de la plupart des cultivars diffusés en Amérique. Pourtant, l'ancêtre sauvage est endémique des basses terres, et absent des hautes terres. Ce paradoxe a été résolu récemment (van Heerwaarden et al. 2011). Sur les hauts plateaux, des flux de gènes avec une troisième sous-espèce de téosinte, *Zea mays* ssp. *mexicana*, génétiquement proche de ssp. *parviglumis*, ont modifié le signal génétique initial, en augmentant « artificiellement » la similarité génétique entre les populations de maïs des hauts plateaux et *Zea mays* ssp. *parviglumis*. Plusieurs parents sauvages ont ici participé au compartiment cultivé actuel.

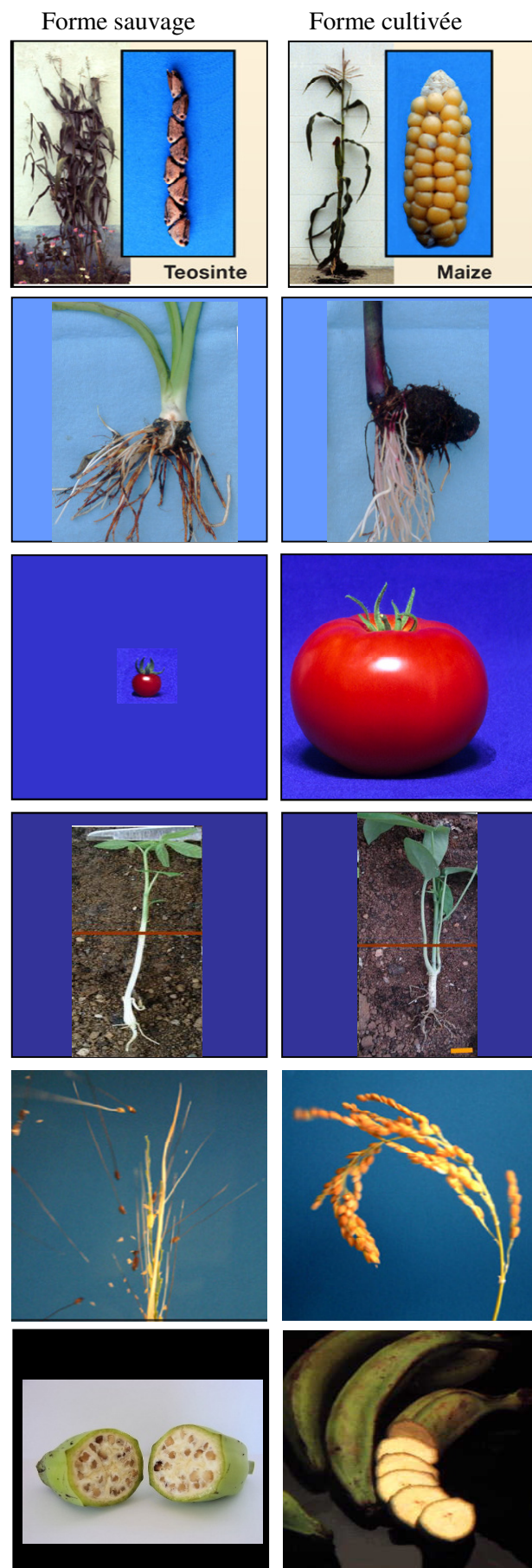
Seul un échantillonnage exhaustif associé à une combinaison de marqueurs appropriés permet donc d'élucider cette question de l'origine (Gepts 2004, Burger et al. 2008). L'utilisation de nouvelles méthodes statistiques, notamment d'inférence bayésienne, exploitant la théorie de la coalescence, permettent aujourd'hui de tester différents types de scénarios démographiques (Kuhner 2009).

### ***1.1.3. Temporalité***

La relative simultanéité avec laquelle apparaissent les différents sites archéologiques attestant des débuts de l'agriculture au travers du monde amène à rechercher des causes communes pour expliquer l'avènement de cette transition (Piperno 2011) (Figure 1-1). Pourquoi les humains durant plus de 95% de leur histoire se sont satisfaits d'une appropriation de ressources spontanées pour leur subsistance et se sont seulement lancés durant la première moitié de l'Holocène dans la production de nourriture ?

Les tenants d'une vision « environnementaliste » évoquent le rôle du climat et/ou de la démographie. Pour certains, la population humaine, de par ses modes de prédation de plus en plus diversifiés, était en pleine expansion à l'émergence de l'agriculture. Elle aurait ainsi pu atteindre en certaines régions les limites d'exploitation de son milieu (Cohen 1977). Pour d'autres, les conditions climatiques prévalent (Richerson et al 2001). Le Pléistocène supérieur est caractérisé par une longue période glaciaire marquée par d'importantes fluctuations climatiques à des échelles de temps courtes, inférieures au millénaire. Le climat y est globalement aride et le taux de CO<sub>2</sub> atmosphérique faible, limitant ainsi la production primaire végétale. Ces conditions environnementales auraient globalement été défavorables à l'installation de l'agriculture, un système d'exploitation des ressources certes efficace, mais particulièrement vulnérable aux variations climatiques. À l'Holocène, le climat se réchauffe et surtout se stabilise. Aussi, la végétation connaît de profondes modifications. Ces transformations auraient pu donner aux hommes l'opportunité d'une telle transition. Même si de telles raisons sont probablement vraisemblables, il ne faut pas céder à la simplification mécaniste et voir dans le climat les seules raisons de la domestication. Certains auteurs y préfèrent notamment des explications socioculturelles et estiment que cette révolution économique a dû être précédée d'une révolution socioculturelle et symbolique (Cauvin 1997). Les datations obtenues à partir des premiers restes archéobotaniques dans les différentes régions du monde attestent que la domestication a débuté dès 10 000 ans dans le Croissant Fertile, entre 9 000 et 8 000 ans en Amérique et en Chine, autour de 7

**Figure 1-2: Phénotypes comparés de formes sauvage et cultivée**



### **Syndrome de domestication**

#### De la téosinte au maïs:

- Compaction de l'architecture de la plante (une tige par plante)
- Augmentation de la taille du caryopse et de l'infrutescence
- Enveloppe protectrice réduite

#### Le taro:

- Augmentation de la taille du corme
- Diminution de la teneur en acide oxalique du corme
- Raccourcissement des stolons

#### La tomate:

- Augmentation significative de la taille du fruit

#### Le manioc

- Augmentation de la taille du tubercule
- Augmentation du diamètre de la tige et de sa teneur en réserves
- Multiplication des méristèmes secondaires
- Modification de la morphologie fonctionnelle de la plantule : d'une germination hypogée à épigée (le trait rouge symbolise le niveau d'enfouissement)

#### Le riz

- Compaction des panicules
- Augmentation de la taille des grains
- Perte de la dispersion à maturité
- Perte de la dormance des graines
- Passage d'un système de reproduction mixte à de l'autofécondation presque stricte.

#### Le bananier

- Parthénocarpie
- Augmentation de la taille du fruit et de la teneur en pulpe



000 ans en Nouvelle-Guinée, et plus récemment en Afrique, autour de 4 500 ans (Purugganan et Fuller 2009, Price et Bar Yosef, 2011). Derrière cette relative simultanéité se dissimule donc une certaine hétérogénéité qui traduit bien le fait que cette transition résulte d'un enchaînement complexe de facteurs environnementaux et socioculturels, qu'il convient d'étudier au cas par cas (Vrydaghs et Denham 2007, Piperno 2011).

La transition de la cueillette vers l'agriculture et l'obtention de plantes domestiquées n'a rien d'un événement. Les archéologues la représentent souvent comme une succession d'étapes (Harris 1989, Fuller 2007) qui comprend en général 1) la cueillette de formes spontanées, 2) la production de plantes sauvages, 3) la mise en culture de formes sauvages puis 4) la culture de plantes domestiquées (agriculture au sens strict). Selon le contexte géographique, ces différentes étapes n'ont pas eu la même durée. De plus, il est à noter que dans certaines régions, comme en Mélanésie par exemple, ces différentes formes d'exploitation de l'environnement végétal se rencontrent toujours aujourd'hui sous la forme d'une mosaïque de pratiques où cueillette et pratiques agricoles co-existent et ne peuvent donc pas être envisagées comme une succession linéaire de séquences temporelles (Barrau 1956, Allaby 2007, Walter et Lebot 2007).

Dans une région géographique donnée, toutes les plantes n'ont d'ailleurs pas été domestiquées à la même période. En Mésopotamie, le maïs a probablement été domestiqué il y a 9 000 ans (Ranere et al. 2009), alors que le haricot l'a été seulement plus tardivement, il y a 4 000 ans (Kaplan et Lynch 1999, Chacon et al. 2005). Certaines espèces n'ont été domestiquées que très tardivement. On peut citer le guarana (*Paullinia cupana*), petit arbuste de la famille des Sapindacées, cultivé en Amazonie pour sa graine, qui aurait été domestiqué il y a moins de 2000 ans (Clement et al. 2010) ; le kava (*Piper methysticum*), domestiqué il y a moins de 3000 ans probablement au Vanouatou (Lebot et al. 1997); ou encore le pacanier (*Carya illinoensis*), arbre produisant la noix de pécan, domestiqué il y a à peine 150 ans dans les plaines alluviales du centre des Etats-Unis (Miller et Gross 2011).

La domestication doit donc être envisagée comme un processus diffus, à la fois dans le temps et dans l'espace, processus qui recouvre des réalités différentes dans les différentes localités où cette transition a été amorcée.

## I.2. Le(s) processus de domestication

La domestication est un processus évolutif à travers lequel une population, soumise à sélection, diverge et s'adapte à de nouvelles conditions environnementales créées par l'homme, certaines de ces adaptations pouvant être délétères en milieu naturel (Gepts 2004). Une plante est dite domestiquée lorsqu'elle présente des caractéristiques (morphologiques, génétiques ou écologiques) la distinguant de son ancêtre sauvage (Purugganan et Fuller 2009). On parle de syndrome de domestication.

### I.2.1. Syndromes de domestication

Gepts (2004) et Pickersgill (2007) proposent une description complète des syndromes de domestication que nous résumons ici. Ce syndrome touche en premier lieu les organes de la plante directement convoités par l'homme comme les fruits, les graines et certains organes végétatifs (racines, tubercules, bulbe, cormes, fruits...) (Figure 1-2). Ces organes sont en général de taille plus importante et présentent une plus forte variabilité morphologique chez le cultivé. Chez les céréales, les deux caractéristiques majeures de domestication sont la perte de la dispersion et de la dormance des graines (Glémin et Bataillon 2009, Fuller 2007). On note aussi des modifications du port de l'appareil végétatif, qui est alors souvent plus compact, moins ramifié, les formes rampantes étant souvent remplacées au profit de formes érigées. On observe aussi la diminution ou perte des mécanismes de défenses naturelles, qu'ils soient chimiques (composés toxiques) ou mécaniques (épines), en particulier s'ils concernent la partie consommée. On note enfin, des changements considérables au niveau des systèmes de reproduction des formes cultivées. Ainsi, de nombreuses plantes ont évolué de l'allogamie (fécondation croisée) vers l'autogamie (autofécondation), comme le riz (Glémin et Bataillon, 2009) et d'autres ont même perdu leur système d'auto-incompatibilité. Certaines ont évolué de la dioécie (les deux sexes portés sur des plantes différentes) vers l'hermaphrodisme (les deux sexes

portés sur la même fleur) comme la vigne, *Vitis vinifera* (This et al. 2006). Enfin, chez certaines plantes, la parthénocarpie ou d'autres formes de stérilité ont finalement été retenues au cours du processus de domestication comme chez le bananier (*Musa* sp.) ou l'arbre à pain (*Artocarpus* sp.).

### ***1.2.2. Forces à l'œuvre***

Cet état « domestique » est généralement considéré comme l'aboutissement d'un processus sélectif, qui débute par l'exploitation préférentielle de certaines formes sauvages présentant des caractères clefs, leur mise en culture dans des espaces aménagés aux contraintes sélectives différentes de celles de leurs écosystèmes naturels, et la fixation de modifications entre formes sauvages et cultivées (Pickersgill 2007). Il y a très peu de chance que l'ensemble des mutations qu'il a fallu assembler pour obtenir un phénotype domestiqué donné soient apparues chez le même individu ou dans la même population (Gepts 2004, Mc Key 2012). C'est donc la succession de cycles de recombinaison-sélection qui a conduit à l'obtention du syndrome de domestication. Dans certains cas, la domestication d'une plante ne nécessite pas de véritable mise en culture, c'est-à-dire d'être entretenue dans un milieu aménagé par l'homme. Certaines pratiques, répandues en Mésio-Amérique (Casas et al. 2007) ou en Océanie (Barrau 1962, Yen 1974, Allaby 2007, Walter et Lebot 2007) notamment pour l'exploitation des arbres, et qualifiée de domestication *in situ* (Casas et al. 1997), consistent à supprimer dans le milieu naturel les phénotypes indésirables et/ou favoriser les phénotypes intéressants. Cette simple sélection artificielle d'individus modifierait progressivement la composition phénotypique puis génétique des peuplements sauvages. Ainsi en Océanie, les Papous des basses terres marécageuses de Nouvelle-Guinée, dans les régions des fleuves Sepik et Fly, vivent de la fécule prélevée dans le stipe du sagoutier, le palmier *Metroxylon sagu*. Ces populations exploitent depuis longtemps des peuplements spontanés de sagoutiers. Par un long processus de sélection empirique, ils ont ainsi éliminé de la presque totalité des palmeraies, les formes épineuses qui se manipulent difficilement (Barrau 1962).

Ces traits ont pu être sélectionnés de manière consciente ou inconsciente (Zohary 2004) : l'augmentation de rendement ou l'amélioration du goût et de la texture des parties consommées a bien dû résulter de la sélection intentionnelle par les agriculteurs de certains individus particulièrement attractifs. Par contre, d'autres adaptations non visibles à l'œil nu, comme la perte de la dispersion chez les céréales, la perte de la dormance, ou encore l'évolution du système de reproduction, ont probablement été sélectionnées de manière inconsciente par l'homme, mais indirectement en réponse aux nouvelles pressions de sélection du milieu cultivé.

En domestiquant, l'homme a prélevé une fraction seulement du compartiment sauvage, créant ainsi un effet de dérive dans ces populations. On parle de goulots d'étranglement. La dérive conduit à la perte aléatoire de certains allèles et donc à une réduction globale de la diversité génétique au sein du compartiment cultivé par rapport au compartiment sauvage. Cette diminution est variable selon la plante considérée. Elle peut atteindre jusqu'à 75% chez le riz (*O. sativa* ssp. *japonica*) (Caicedo et al. 2007) ou encore 60% chez le tournesol (Liu et Burke 2006) et seulement 30 % pour le mil (*Pennisetum glaucum*) (Oumar et al. 2008) ou 40% chez le maïs (Wright et al. 2005). Le processus de sélection a lui aussi entraîné une perte de diversité mais cette fois accentuée au niveau de régions ciblées du génome, qui contrôlent l'expression des traits de domestication. On parle de balayage sélectif. De ce processus d'échantillonnage et de sélection ciblée résulte une divergence entre populations cultivées et populations sauvages. L'importance des flux de gènes entre ces deux compartiments a ensuite largement conditionné la rapidité et le rythme même du processus de domestication. Bien que de nouvelles appellations taxonomiques soient souvent attribuées aux formes cultivées, la domestication n'a conduit que rarement à la divergence d'une nouvelle espèce (au sens biologique, c'est-à-dire reproductivement isolée de son ancêtre sauvage) (Harlan 1992).

Derrière l'image d'une relative convergence évolutive des syndromes de domestication, il existe en fait une large diversité de combinaison de traits « domestiqués », ainsi que différents niveaux de domesticité.

Comment domestique-t-on une plante ? Cette question a probablement autant de réponses qu'il existe de plantes domestiquées. La diversité des syndromes révèle en fait la multitude de trajectoires évolutives qui conduisent à l'obtention d'une plante domestiquée, trajectoires qui

dépendent de la biologie de la plante, de l'environnement de domestication, et du contexte culturel associé. Nous allons tâcher ici d'en illustrer quelques aspects.

### ***1.2.3. Domestication des plantes annuelles propagées par graines : l'exemple des céréales***

Les céréales sont des plantes annuelles, propagées par graines et appartenant à la famille des *Poaceae*. Elles sont caractérisées par un syndrome de domestication marqué et relativement homogène (Figure 1-2), qui suggère l'existence d'une convergence évolutive des processus de domestication, certains auteurs proposant même que cette évolution parallèle résulte de l'évolution indépendante de traits domestiqués à partir de gènes orthologues (Paterson et al. 1995). Ce syndrome inclue en général les six traits suivants : élimination/ réduction de la capacité de dispersion, réduction des structures d'aides à la dispersion, augmentation de la taille des graines, perte de la dormance, synchronisation du développement et compaction du mode de croissance (réduction des ramifications axillaires) (Glemin et Bataillon 2009). Dans le Croissant Fertile, les céréales sauvages ont été cultivées bien avant leur domestication. Les premières formes domestiquées de céréales apparaissent il y a 10 500 ans, et dans le même temps, les formes sauvages persistent, dans des proportions non négligeables, pendant au moins un millénaire (Willcox et Tanno 2006).

#### *Bases génétiques de la domestication.*

L'analyse des bases génétiques du syndrome de domestication a été principalement menée sur les plantes annuelles, dont les céréales, et a révélé plusieurs caractéristiques intéressantes.

La cartographie génétique des loci impliqués dans la variation des caractères clés de la domestication a montré que le syndrome de domestication est majoritairement contrôlé par un nombre restreint de loci qui ont des effets majeurs. De plus, ces loci sont souvent regroupés dans le génome (Gepts 2004, Burger et al. 2008, Glemin et Bataillon 2009). Cette architecture génétique aurait ainsi pu permettre une fixation relativement rapide des traits de domestication, limitant le nombre de cycles de recombinaison-sélection nécessaires. Alternativement, le processus de domestication lui-même aurait pu sélectionner la liaison physique de ces gènes. Des études plus récentes, recherchant un signal de sélection cette fois-ci sur l'ensemble du génome, suggèrent toutefois qu'un plus grand nombre de gènes pourraient finalement être impliqués (plus de 1200 gènes pour le maïs, par exemple) (Wright et al. 2005, Caicedo et al. 2007).

Les loci impliqués dans un même trait du syndrome de domestication peuvent aussi différer entre deux espèces. Par exemple, le gène *tb1* largement impliqué dans les changements architecturaux entre le maïs et la téosinte (Doebley et al. 1997) semble peu affecter l'architecture du petit mil (*Setaria italica*) (Doust et al. 2004 et 2007). Enfin, au sein d'une espèce domestiquée, un même phénotype peut être contrôlé par différentes mutations. C'est par exemple le cas du gène *vrs1* chez l'orge où trois mutations indépendantes contrôlent l'apparition d'un même phénotype (Komatsuda 2007), ou encore le gène *Rc* du riz, où deux mutations différentes contrôlent la perte de pigmentation du péricarpe chez les formes cultivées (Mc Couch et al. 2012).

Ainsi l'architecture génétique des traits de domestication suggère que la domestication a pu être relativement rapide (Gepts 2004). Néanmoins, des voies génétiques différentes ont conduit à l'obtention de phénotypes domestiqués semblables pour les différentes espèces, ainsi qu'au sein d'une même espèce.

Des expériences de reconstitution de domestication menées sur le blé et l'orge, ainsi que des simulations mathématiques ont pu suggérer qu'un caractère comme la perte de la dispersion avait évolué rapidement (Hillman and Davies 1990). La domestication était donc appréhendée comme un processus relativement rapide, avec des plantes domestiquées à partir d'un nombre restreint d'individus provenant d'une aire géographique limitée (Hillman and Davies 1990, Honne and Heun 2009). Ce processus aurait pu être d'autant plus rapide que la plupart des céréales, à l'exception de deux exemples notables (le maïs et le mil), sont des plantes majoritairement autogames, qui peuvent donc théoriquement fixer rapidement les « allèles de domestication » (Glémin et Bataillon 2009). Les premières données génétiques, basées sur des marqueurs neutres comme les AFLP ou les microsatellites attestaient d'ailleurs de l'origine unique de la plupart des céréales corroborant ainsi ce scénario (ex. Heun et al. 1998 pour l'engrain -*Triticum monococcum*- ; Badr et al. 2000 pour l'orge, -*Hordeum vulgare*-).

Depuis, les archéologues ont reconstruit l'évolution temporelle de deux traits majeurs du syndrome de domestication des céréales, la taille du grain et la perte de la capacité de dispersion (Fuller 2007 et 2012, Purugganan et Fuller 2009). L'analyse des restes archéobotaniques du blé et de l'orge dans le Croissant Fertile montre que l'augmentation de la taille des grains, caractère facilement sélectionnable par l'homme, a été progressive mais achevée rapidement, en moins de 500 ans. Par contre, l'évolution de la perte de la dispersion a été beaucoup plus lente et aurait mis environ 3000 ans à être fixée. L'étude des restes archéobotaniques obtenus pour le riz en Asie dessinent un scénario similaire (Fuller 2007 et 2012).

Ces données témoignent d'un processus de domestication graduel, les différents traits se fixant les uns après les autres, et prolongé, plus de 3000 ans de mise en culture avant la fixation de certains traits. Parallèlement, les données génétiques obtenues récemment sous-tendent une origine géographique multiple et/ou l'existence de phénomènes d'introgession prolongée entre populations sauvages et populations cultivées pour ces trois céréales (ex. Mc Couch et al. 2012 pour le riz, Morrell et al. 2007 pour l'orge, Killian et al. 2007 pour l'engrain). La domestication se serait ainsi déroulée sur une période longue, des populations auraient été domestiquées dans différentes localités en sympatrie avec les populations sauvages, et auraient entrete nu entre elles des flux de gènes importants d'un lieu à un autre, ralentissant la fixation du syndrome de domestication (Allaby et al. 2008, Purugganan et Fuller 2009).

#### ***1.2.4. Domestiquer et reproduire une plante par propagation clonale: l'exemple des plantes à racines et tubercules tropicales***

Alors que certaines plantes ont été domestiquées puis reproduites à partir de graines, d'autres sont propagées clonalement par les agriculteurs. Ainsi, chaque plante de la nouvelle génération est une copie génétique identique - aux mutations somatiques près - d'une plante déjà présente dans la génération précédente. Cela concerne des plantes réparties dans de nombreux taxons phylogénétiques avec plus de 34 familles concernées, et représentant des types biologiques très variées autant sur le plan écologique que morphologique (Mc Key et al. 2010 et 2012). Elles sont cultivées pour leurs fruits, des organes de stockage sous-terrain (racines, rhizomes, tubercules, cormes...), leurs feuilles et même parfois leurs graines. Les parties du végétal utilisées pour le bouturage sont aussi très variées (morceau de tige, de tubercules, rejets...). On distingue les plantes pérennes, des plantes à multiplication végétative annuelle (racines ou tubercules par exemple). Les plantes pérennes sont en général caractérisées par un syndrome de domestication faible et diffus et une faible réduction de la diversité génétique entre les populations sauvages et cultivées. Cela reflète le faible nombre de cycles sexués séparant les individus domestiqués de leurs ancêtres sauvages, en lien avec la longue phase juvénile de ces plantes pérennes, ainsi que l'importance des flux de gènes entre les populations sauvages et cultivées (Miller et Gross 2011).

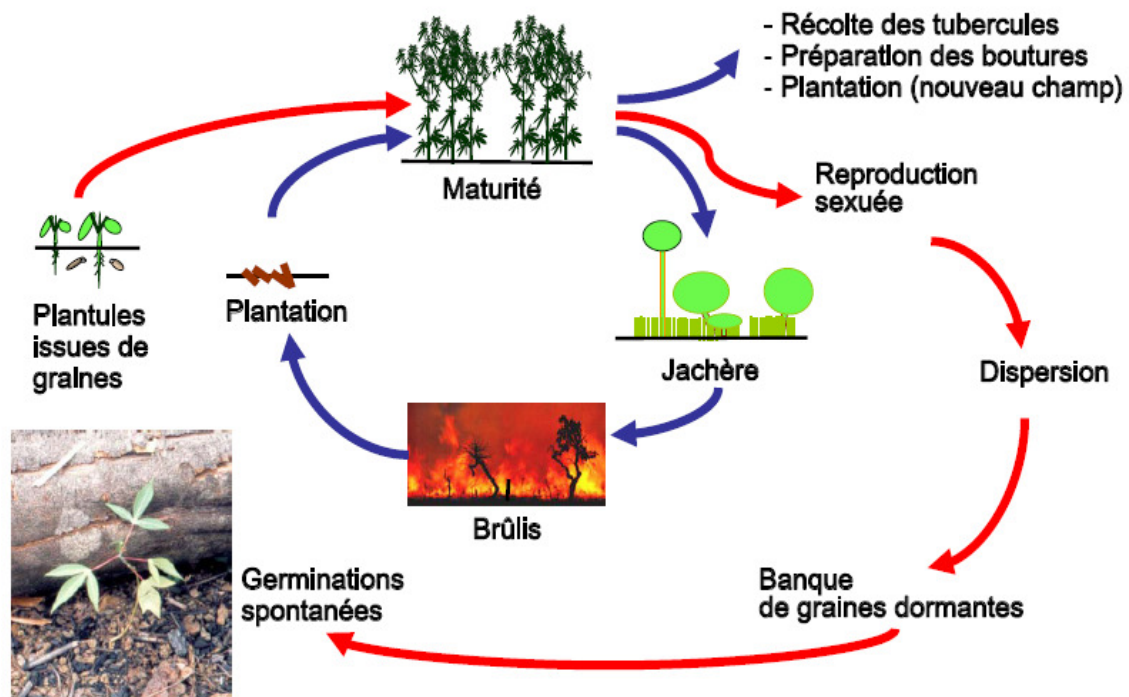
Nous allons nous intéresser ici plus particulièrement au cas des racines et tubercules, plantes à multiplication végétative annuelle.

Dans certains cas, la propagation clonale a tout simplement pu être choisie comme moyen le plus simple de propager la plante, un certain nombre d'espèces sauvages possédant déjà des structures qui permettent la propagation clonale, notamment chez les tubéreuses tropicales (Mc Key et al. 2010). Un deuxième avantage est que la bouture contient souvent plus de « réserves » qu'une graine, ce qui permet un développement et une croissance plus rapides de la plante et une meilleure survie. Enfin, la plupart des plantes multipliées végétativement se trouvent être des allogames, plus ou moins obligatoires, c'est-à-dire qu'elles possèdent des systèmes de reproduction qui favorisent et même obligent les croisements entre individus différents (Zohary 2004, Mc Key 2010 et 2012). Leurs descendants expriment donc souvent des phénotypes contrastés, différents de ceux des parents, ralentissant ainsi la fixation de traits d'intérêt au fil des générations. La propagation clonale apparaît donc comme un moyen immédiat de fixer puis de reproduire un caractère d'intérêt.

Ainsi des mutants sauvages présentant des caractères attractifs pour le cultivateur auraient pu être rapidement capturés puis reproduits à l'identique en milieu cultivé. Certains auteurs ont parlé de domestication immédiate (Zohary 2004). Ce scénario implique que très peu de cycles de recombinaison-sélection séparent la forme cultivée de la forme sauvage, qui doivent donc différer relativement peu l'une de l'autre. Il existe une plante pour laquelle un tel scénario de domestication instantanée est envisageable, il s'agit du kava, domestiqué probablement au Vanouatou il y a moins de 3 000 ans (Lebot et al. 1997) à partir de l'espèce sauvage (conspécifique) *P. wichmanii*. Les différents cultivars de kava que l'on trouve aujourd'hui dans le Pacifique ont une base génétique très étroite et pourraient tous dériver d'un seul individu par pure sélection clonale, les différents cultivars rencontrés aujourd'hui ayant été l'objet d'une sélection rigoureuse de mutants somatiques par les cultivateurs. A moins que les autres clones domestiqués n'aient pas survécu.

Pourtant, ce scénario de domestication ne semble pas tenir pour d'autres espèces multipliées végétativement. Il est d'abord rare que la base génétique des cultivars d'une espèce à propagation clonale soit extrêmement étroite (ex. pomme de terre -*Solanum tuberosum*- : Brush et al. 1995 ; manioc : Elias et al. 2001 et 2004, Sardos et al. 2008, Duputié 2009 ; taro : Lebot et al. 2004, Caillon et al. 2006, ignames -*Dioscorea* sp.- : Malapa et al. 2005, Tostain et al. 2007), hormis à une échelle locale. En général, elle est constituée de groupes de génotypes distants les uns des autres, souvent très hétérozygotes (Elias et al. 2004, Fregene et al. 2003, Duputié et al. 2009), attestant de l'existence d'événements de recombinaison sexuée à l'origine des différents clones sélectionnés. Aussi, dans certains cas, les syndromes de domestication de ces plantes englobent plusieurs traits complexes qui ont peu de chance d'avoir été assemblés et stabilisés au sein du compartiment cultivé par une stricte sélection clonale, sur une période de temps aussi courte que 10 000 ans (Mc Key et al. 2010 et 2012). A ce titre, l'exemple du manioc est assez éloquent et a fait l'objet de nombreux travaux (revue par Rival et Mc Key 2008, Mc Key et al. 2010 et 2012). Le syndrome de domestication du manioc inclue des traits directement sélectionnables par les agriculteurs, comme l'augmentation du rendement des tubercules. D'autres traits confèrent une meilleure aptitude à la propagation clonale, comme l'augmentation du diamètre de la tige qui permet une meilleure conservation de la bouture, l'augmentation de sa teneur en réserves, ou encore la multiplication des méristèmes secondaires qui garantissent un meilleur taux de reprise végétative. D'autres adaptations ont probablement été générées par la sélection du milieu cultivé. Parmi elles, la plus frappante concerne la morphologie fonctionnelle de la plantule (Figure 1-2). Les graines de manioc sauvages ont une germination hypogée, où les cotylédons restent enfouis dans le sol (l'hypocotyle ne croissant pas en longueur) et enclos dans le testa et servent de réserves. Cela confère à la plante une croissance initiale lente, mais qui lui permet de garder une grande tolérance à la perte des parties aériennes, par l'existence de réserves (cotylédons plus hypocotyle) et de méristèmes (axillaires aux cotylédons et sur la partie souterraine de l'épicotyle) dans le sol. Lorsque les graines sont enfouies très peu profondément, les cotylédons peuvent parfois être portés par des pétioles hors du testa et devenir foliacés, cependant, l'allongement de l'hypocotyle reste toujours faible. Ainsi le manioc sauvage est adapté à des milieux imprévisibles et pauvres en ressources, où les risques de perte des parties aériennes sont importants. En revanche, chez le manioc domestiqué, on observe une germination épigée. Cela implique la mise en place très rapide de cotylédons photosynthétiques, qui offrent à la plante une croissance initiale

**Figure 1-3 : Cycle de vie du manioc domestiqué dans les systèmes de culture amérindiens.**  
 La complémentarité des sous-cycles clonal (en bleu) et sexué (en rouge) est présentée. (D'après Mc Key et al. 2012).



rapide, probablement avantageuse en milieu cultivé, mais défavorable en milieu naturel, puisque la plante est très sensible à la perte de ses parties aériennes, ne possédant ni réserves, ni méristèmes souterrains (Pujol et al. 2005).

Une telle évolution est difficile à entrevoir par propagation clonale étant donné les limites évolutives d'une telle propagation (Barton et Charleworth 1998, Mc Key et al. 2010 et 2012). Les clones ne sont pas immuables et peuvent accumuler des mutations. Certaines de ces mutations sont neutres, n'ayant pas d'incidence sur la valeur sélective de la plante ; d'autres ont un effet délétère, d'autres encore ont un effet positif. Les mutations ayant un effet négatif semblent être beaucoup plus fréquentes que les mutations bénéfiques (Fisher 1930). La plante accumulerait donc de plus en plus de mutations délétères avec le temps, processus dénommé « cliquet de Muller » qui pourrait en théorie conduire à la diminution des performances et de la qualité agronomique du clone avec le temps (Mc Key et al. 2010). De plus, lorsque la propagation est strictement clonale, il n'y a aucun mécanisme qui permette d'associer les différentes mutations bénéfiques qui ont pu se produire dans différentes lignées clonales. La recombinaison élimine la plupart de ces mutations délétères et combine les différentes mutations favorables au sein d'un même individu. Elle permet ainsi une évolution plus rapide que la reproduction strictement clonale (Mc Key et al. 2010).

Même si certaines plantes à propagation clonale présentent aujourd'hui une fertilité très réduite, voire même absente (le kava, le bananier, ou encore certaines espèces d'ignames par exemple), beaucoup d'entre elles ont gardé leur capacité à se reproduire sexuellement. Cette reproduction sexuée a probablement joué un rôle primordial dans l'histoire évolutive précoce de ces plantes, particulièrement celle de leur domestication. Les restes archéobotaniques des plantes à racines et tubercules ne permettent pas de suivre l'évolution des traits de domestication comme dans le cas des céréales. Par contre, l'observation des pratiques actuelles des cultivateurs offre des pistes intéressantes pour comprendre le processus de domestication et en particulier les modalités de coexistence, pour ces espèces, d'un régime clonal et d'un régime sexué (Mc Key 2012). Deux exemples contrastés ont été bien documentés, le manioc en Amérique du Sud et l'igname en Afrique de l'Ouest.

#### *L'exemple du manioc dans les systèmes agricoles amazoniens*

Dans les systèmes d'agriculture itinérante sur brûlis, le manioc est généralement reproduit végétativement par bouturage des morceaux de tiges. La racine tubérisée du manioc se conservant mal, les pieds de manioc sont en général laissés en place dans les champs et ont le temps de fleurir et de produire des graines, qui tombent au sol (Elias et al. 2000 et 2001). Ces graines sont ensuite déplacées par des fourmis qui les entassent à relative proximité de la surface du sol. Elles restent dormantes, attendant une augmentation significative de la température du sol (qui signale une ouverture du milieu), pour germer. Le jardin est alors laissé en jachère entre 10 et 25 ans. Plus tard, ces friches sont de nouveau ouvertes (par brûlis, ou plus simplement par défrichage), générant des conditions favorables à la germination des graines dormantes (Pujol et al. 2002 et 2007). Ainsi, lorsque ces nouveaux jardins sont cultivés, de jeunes plantules de manioc issues de graines font surface. Les agriculteurs en conservent alors une partie dans leurs champs, qu'ils intègrent plus tard à leur stock de variétés (Figure 1-3). Seulement, l'intégration de ces recombinants n'est pas le simple fruit du hasard, mais le résultat d'un processus sélectif complexe où agissent de concert sélection consciente des cultivateurs et sélection inconsciente exercée par le milieu directement, favorisant toutes deux la reproduction de plantules vigoureuses (Pujol et al. 2005), qui croissent vite et correspondent aussi à un type donné (Duputié et al. 2009). D'abord, lors du désherbage manuel de la parcelle (Pujol et al. 2005), les agriculteurs retireraient systématiquement les petites plantules, qui sont aussi les plus homozygotes, ne les repérant pas (sélection humaine inconsciente et /ou consciente), alors qu'ils laisseraient les plus vigoureuses. Puis, la mortalité naturelle viendrait achever la deuxième étape de sélection (Pujol et McKey, 2006), mortalité qui serait due à la compétition entre plantes issues de graines, souvent agrégées dans les poubelles à fourmis. Les plantes les plus grandes, qui sont aussi les plus hétérozygotes, survivraient préférentiellement. Dans la troisième étape de sélection, les agriculteurs, lors de la récolte, sélectionneraient les plantes issues de graines pour la préparation des boutures et

favoriseraient là encore les plus vigoureuses. Enfin, ces clones issus de graines seraient testés quelques années avant adoption définitive et reproduction par voie clonale, les clones hétérozygotes continuant encore à être favorisés lors de cette dernière étape de sélection (Duputié et al. 2009). Dans le même temps, il semble que les agriculteurs ne se contentent pas de contre-sélectionner les plantes les plus homozygotes, mais aussi celles qui s'éloignent trop d'un type recherché, pratiquant ainsi une sélection idéotypique (selon un type) (Duputié et al. 2009). Ces pratiques paysannes actuelles pourraient bien expliquer l'évolution de la morphologie de la plantule entre le sauvage et le cultivé. Les plantules présentant une germination épigée, plus attractives pour les agriculteurs à cause de leur croissance initiale plus rapide, auraient ainsi été intégrées préférentiellement (Mc Key et al. 2012).

### *L'igname africaine*

Les ignames du complexe *D. cayennensis/rotundata* sont majoritairement cultivées en Afrique de l'Ouest. Aux alentours des champs poussent souvent ses ancêtres sauvages, *D. abyssinica*, une espèce annuelle qui se rencontre dans des savanes arborées d'Afrique de l'Ouest, et *D. praezensilis*, espèce très proche habitant des forêts denses et humides. Les paysans pratiquent dans ces régions l'ennoblissement, qui consiste à replanter dans les champs des tubercules d'ignames sauvages prélevés dans les savanes ou forêts alentours. Les agriculteurs exercent alors différents stress dans le but de transformer la morphologie de la plante et de l'adapter aux contraintes du milieu cultivé. Ces modifications morphologiques seraient probablement liées à des phénomènes de plasticité phénotypique et des effets épigénétiques (Scarcelli 2005). Après plusieurs années, l'agriculteur estime si l'igname est suffisamment adaptée à la culture et décide ou non de la garder. Cet exemple donne à première vue l'image d'une domestication instantanée des formes sauvages. Pourtant, ces ignames sauvages ennoblies ne sont pas réellement sauvages. En effet, plus la plante ressemble morphologiquement aux ignames cultivées plus elle a de chances d'être sélectionnée (décrit dans Scarcelli 2005). Surtout, une analyse génétique a montré que ces plantes prélevées dans le compartiment sauvage correspondent pour la plupart à des hybrides cultivés-sauvages ou encore à des hybrides inter-variétaux (Scarcelli et al. 2006). Dans cet exemple, les plantes issues de reproduction sexuée passent les premières années de leur vie en milieu naturel et sont donc soumises à d'autres pressions de sélection que les plantules de manioc, qui grandissent en milieu cultivé. Mais là encore, l'alternance de cycles de recombinaison-sélection dans un système mixte clonal-sexué permet d'obtenir des ignames particulièrement aptes à la mise en culture.

Ces deux exemples démontrent l'importance et la complémentarité des compartiments sexués et clonaux dans la dynamique évolutive d'une plante à propagation clonale, dynamique dans laquelle sélection humaine et sélection naturelle ont agi de concert pour la phase de domestication et l'assemblage des traits du syndrome de domestication (Mc Key 2012). Ce système mixte continue d'être tout aussi important pour la diversification de ces plantes à multiplication végétative, permettant à la fois de maintenir le potentiel adaptatif de ces plantes à travers la recombinaison et de fixer rapidement des recombinants attractifs par propagation clonale. Toutefois, au terme de ces deux exemples, il faut bien entrevoir la diversité des trajectoires de domestication possibles pour ces plantes à propagation clonale, diversité qui dépend des modalités de gestion du système mixte et des régimes de pressions de sélection qui se sont exercées sur ces deux compartiments, clonal *versus* sexué (Mc Key 2012).

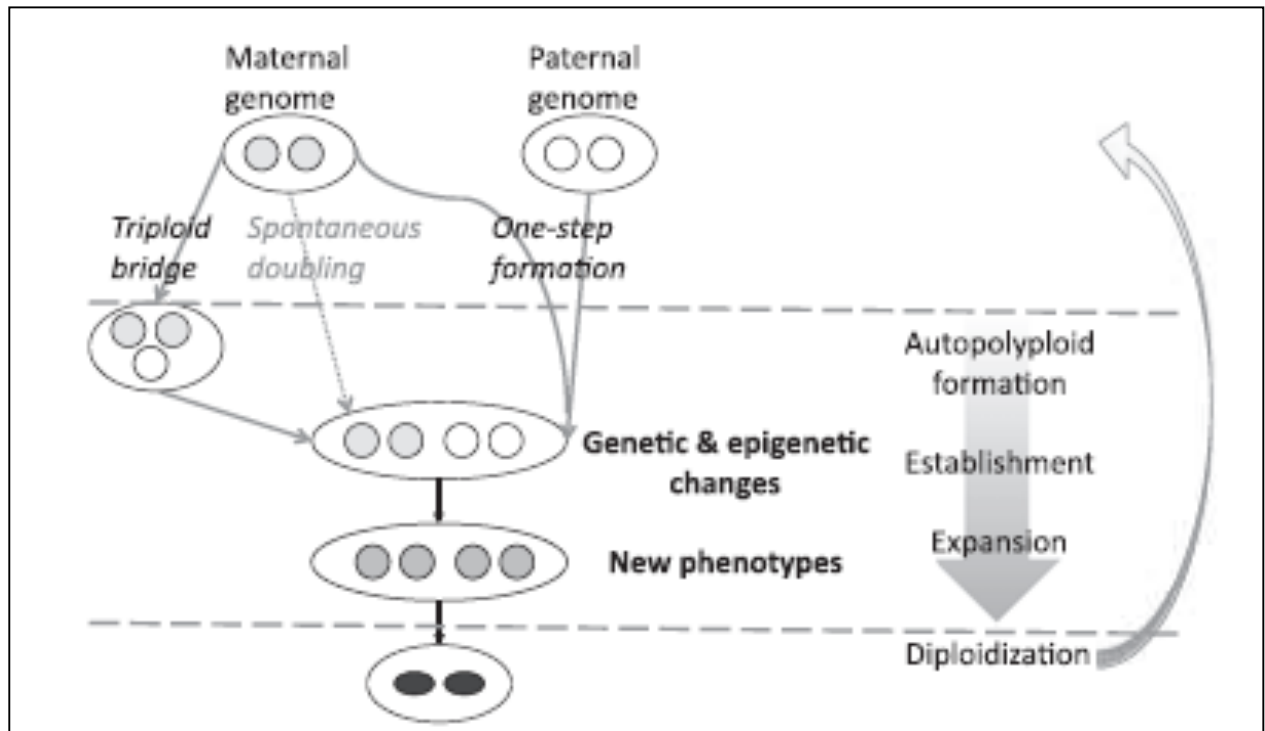
## **1.3. Polyploïdie et domestication**

De nombreuses plantes cultivées sont polyploïdes, c'est-à-dire qu'elles possèdent plusieurs jeux de chromosomes dans un même noyau. Toutefois, la fréquence des polyploïdes chez les plantes cultivées ne diffère pas significativement de celle rencontrée chez les homologues sauvages pour une même catégorie taxonomique (Hilu 1993). Ces résultats semblent indiquer que la polyploïdie ne confère pas d'avantages (tout comme de désavantages) sélectifs particuliers pour le processus de domestication, en



**Figure 1-4 : Voies de formation d'un autopolyploïde et mode de ségrégation polysomique associé (cas d'un autotétraploïde).** (D'après Parisod 2010).

- a) La formation d'un génome autopolyploïde peut résulter de i) la fusion directe de deux gamètes non réduits (« *one-step formation* »), ou bien de deux étapes avec la formation d'un intermédiaire triploïde issu de la fusion d'un gamète non réduit et d'un gamète normal (« *triploid bridge* »). Le doublement spontané somatique, c'est-à-dire sans croisement, (« *spontaneous doubling* ») est rare dans les populations naturelles. Cet état polyploïde donne lieu à des modifications génétiques et épigénétiques qui conduisent à des modifications génomiques d'ordre structural et fonctionnel, à l'origine de nouveaux phénotypes.



tout cas pas plus qu'elle n'en confère en milieu naturel (Hilu 1993). La polyploïdie est considérée comme un facteur majeur dans l'évolution des végétaux notamment des plantes à fleurs (Adams et Wendel 2005). De nombreux taxons ont connu à un moment donné de leur histoire évolutive une phase de polyploïdisation, nommée paleopolyploïdisation, qui a pu s'estomper par la suite au travers d'un processus de re-diploïdisation. On estime même que l'ensemble des angiospermes pourrait ainsi avoir une origine polyploïde (Cui et al. 2006). A long terme, la polyploïdie serait un facteur essentiel de spéciation (Wood et al. 2009), notamment sympatrique (Robertson et al. 2010). De plus, la polyploïdie est souvent considérée comme un élément clé dans la diversification des taxons (Doyle et al. 2008). Une étude récente montre pourtant que si la polyploïdie a un rôle majeur dans la cladogénèse et donc dans la spéciation - on estime aujourd'hui qu'entre 2 et 15 % des événements de spéciation des angiospermes seraient liés à un événement de polyploïdisation -, son implication dans la diversification des taxons est moins claire (Wood et al. 2009).

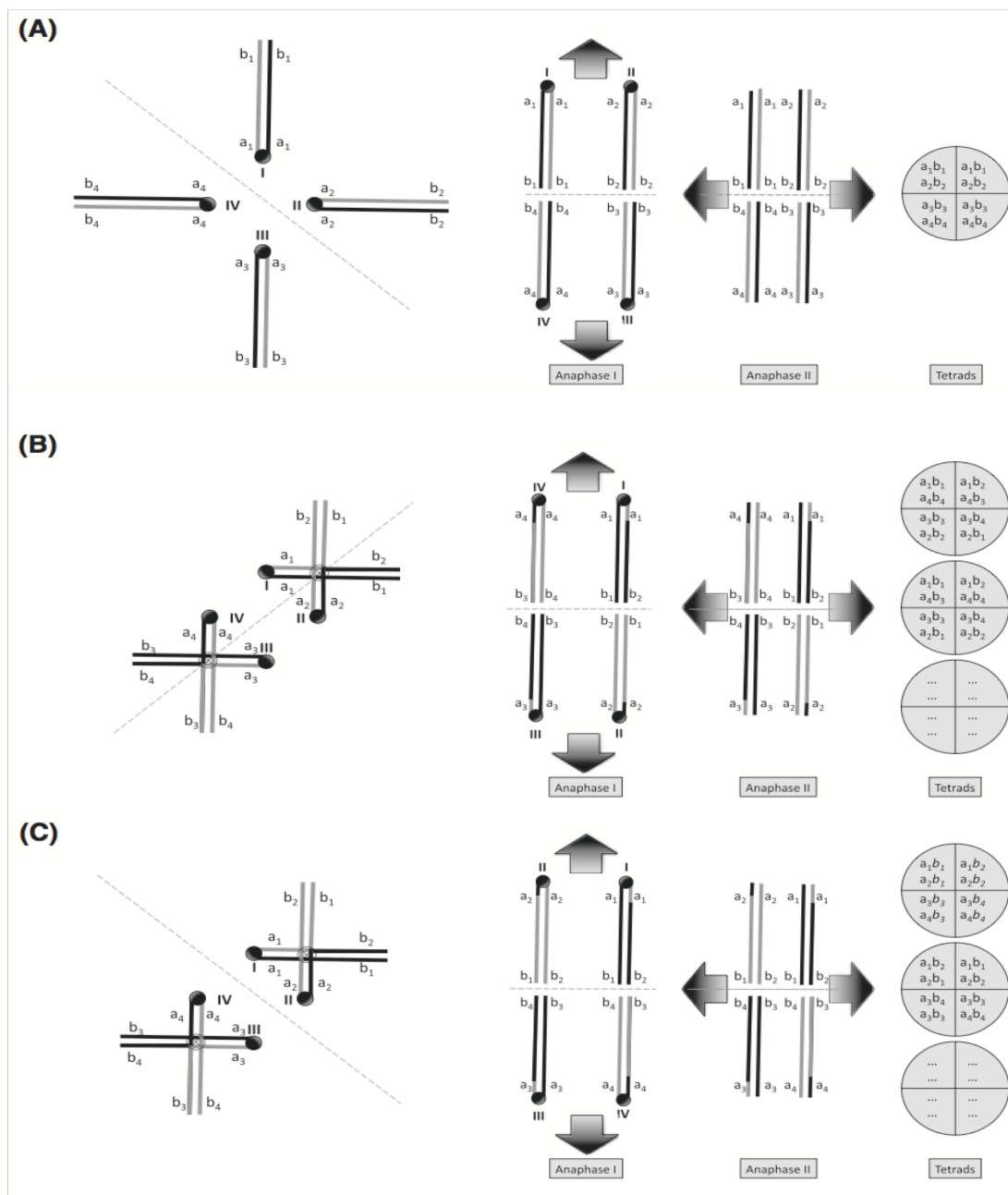
Certains auteurs ont émis l'idée que les plantes polyploïdes présenteraient certaines « prédispositions » à la domestication (Stebbins 1971). Il est remarquable que pour une plante cultivée donnée, lorsque plusieurs niveaux de ploïdie ont coexisté au cours de l'histoire évolutive de la plante, les formes polyploïdes ont bien souvent pris le pas sur les autres. Dans presque toutes les régions où l'engrain ( $2x$ ) et l'amidonnier (*T. turgidum ssp. dicocum*) ( $4x$ ) ont été cultivés ensemble, l'amidonnier est devenu la céréale prédominante. De même, lorsque la forme hexaploïde du blé tendre (*T. aestivum*) apparaît, c'est elle qui sera principalement diffusée (Dubcovsky et al. 2007). Aussi pour la banane, bien que la première forme domestiquée soit diploïde, ceux sont les cultivars triploïdes qui ont gagné en importance et ont été répandus. Quelles sont donc les conséquences évolutives à court terme (puisque le temps d'évolution des plantes cultivées est relativement court, moins de 10 000 ans) qui permettent d'expliquer le succès des plantes cultivées polyploïdes ?

### ***1.3.1. Formation des polyploïdes***

On distingue classiquement deux modes de formations d'une espèce polyploïde. Les autopolyploïdes sont composés de plusieurs jeux de chromosomes homologues issues de la même espèce (Figure 4-1a). Généralement, ils sont caractérisés par une ségrégation dite polysomique, où les différents jeux de chromosomes peuvent s'apparier aléatoirement créant ainsi de nombreuses possibilités de combinaisons alléliques (Parisod et al. 2010) (Figure 1-4b). Chez les allopolyploïdes, les jeux de chromosomes dupliqués proviennent d'espèces différentes (génomes homéologues) mais suffisamment proches pour s'hybrider. En général, ils présentent une ségrégation disomique, où les chromosomes homologues s'apparient préférentiellement. L'autopolyploïdie peut néanmoins impliquer des populations différenciées mais conspécifiques, associant alors des génomes présentant un certain degré de différenciation. Il existe en réalité un continuum qui va de la duplication de génomes strictement identiques (homologues) à celle de génomes hautement différenciés (homéologues) (Parisod et al. 2010). La formation d'un polyploïde en milieu naturel est fréquemment liée à la formation de gamètes non réduites ( $2n$ ), c'est-à-dire l'obtention de gamètes diploïdes issus d'une division meiotique anormale, phénomène relativement courant chez les angiospermes (Otto 2007). Plus rarement, elle intervient par doublement somatique lors du développement du zygote. La polyploïdisation est particulièrement fréquente dans les taxons hybrides (Paun et al. 2008). En effet, l'association de deux génomes très différenciés suite à un événement d'hybridation est instable et conduit à d'importants dérèglements méiotiques, favorisant d'autant plus la formation de gamètes non réduites et l'établissement de lignées polyploïdes dans ces taxons hybrides. Les allopolyploïdes seraient plus fréquents que les autopolyploïdes (autour de 75% des polyploïdes). Aussi, la formation de lignées polyploïdes peut résulter d'un croisement unique, mais plus souvent de croisements multiples et récurrents, limitant dès lors le goulot d'étranglement (réduction de la diversité) associé au processus de polyploïdisation.

**Figure 1-4 (suite) :**

b) La ségrégation polysomique chez un autotétraploïde conduit à la formation de tétravalents en métaphase 1. Au sein des tétravalents, les chromosomes homologues peuvent échanger du matériel génétique par recombinaison si bien que les gamètes produits héritent à chaque locus de deux allèles tirés aléatoirement parmi les quatre allèles possibles. Les patrons de ségrégation et donc la fréquence du type de gamètes obtenue va dépendre de la distance entre le loci impliqués et le centromère et donc des événements de recombinaison potentiels. A) S'il n'y a pas de recombinaison, toutes les combinaisons sont possibles et des chromatides sœurs ne peuvent jamais ségréger dans le même gamète (ségrégation indépendante des chromosomes, première division réductionnelle). S'il y a recombinaison entre le centromère et le locus, les chromatides sœurs peuvent ségréger indépendamment (ségrégation indépendante des chromatides). L'anaphase 1 permet de séparer les 4 chromatides (division 1 équationnelle, B) ou seulement de ségréger des paires de chromatides (division 1 réductionnelle, C). Dans ce dernier cas, l'anaphase 2 permet une double réduction (chromatides sœurs au sein du même gamète). Ainsi, on obtient à la fois des homozygotes et différents types d'hétérozygotes. La plupart des auto-polypléïdes présentent en fait des patrons intermédiaires entre ségrégation disomique et purement tétrasomique, compliquant encore la prédiction des combinaisons obtenues.



### *1.3.2. Avantages évolutifs de la polyploidie dans le contexte de l'évolution des plantes cultivées*

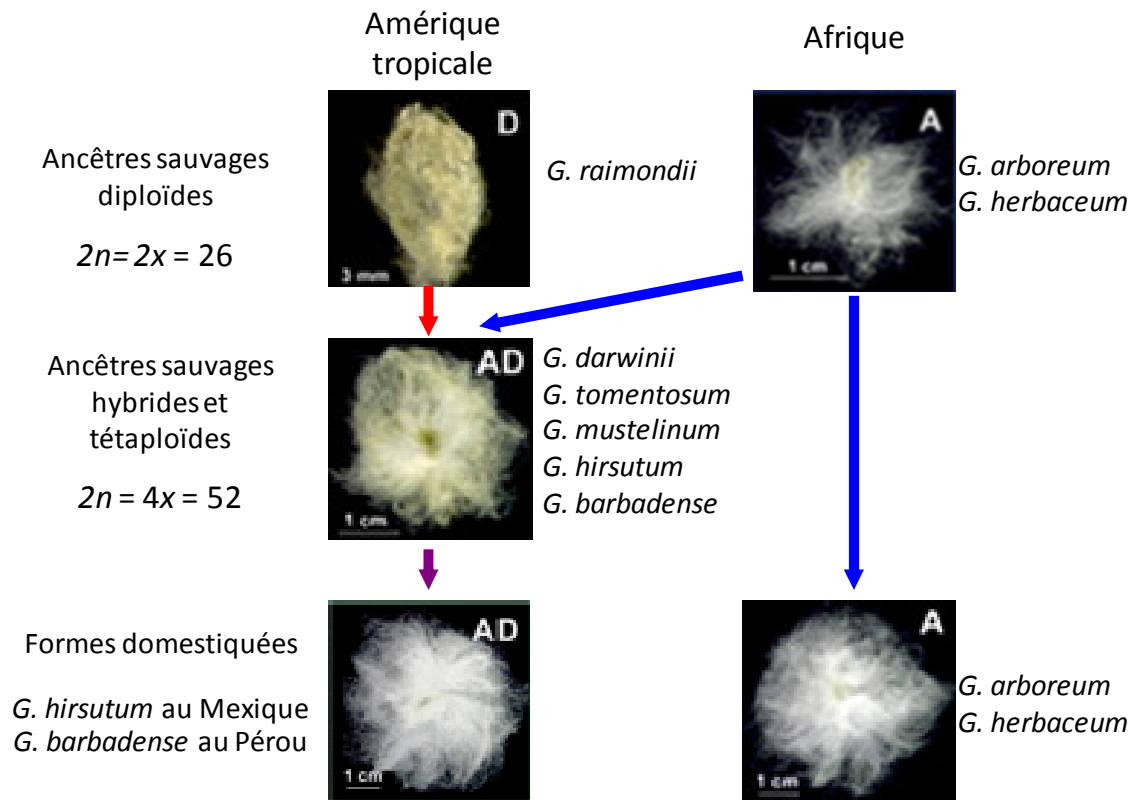
Un événement de polyploïdisation a souvent des effets directs sur le phénotype, liés en premier lieu à l'augmentation de la taille des cellules. Cela peut se traduire par un effet de gigantisme de certains organes (Otto 2007). De plus, les polyploïdes présentent très souvent une augmentation de l'hétérozygotie par rapport à leurs progéniteurs diploïdes et montrent ainsi une vigueur accrue, et ce aux différents stades du développement de la plante (Chen et al. 2010). Dans le cas des allopolyploïdes, cette hétérozygotie est fixée au fil des générations par l'assemblage de génomes différents qui ne peuvent pas recombiner entre eux, on parle d'hétérozygotie génomique. Par contre, chez les autopolyploïdes, elle est liée au mode de ségrégation polysomique et peut évoluer d'une génération à l'autre (Hancock 2004). Nous venons de voir que les agriculteurs de manioc sélectionnent les clones les plus hétérozygotes et aussi les plus vigoureux pour la propagation clonale. Il se pourrait donc que ces plantes particulièrement vigoureuses aient attirées l'œil des cultivateurs, qui les auraient préférentiellement sélectionnées puis reproduites. De plus, cet état hétérozygote permet d'obtenir par recombinaison génétique une grande diversité de combinaisons alléliques (plus grande qu'à l'état diploïde), élargissant ainsi les possibilités de sélection des cultivateurs.

Suite à un événement de polyploïdisation, on observe en général un certain nombre de réarrangements structuraux du génome (disparition et apparition de fragments génomiques, activation d'éléments transposables, modifications des patrons de méthylation de l'ADN), et ce particulièrement chez les allopolyploïdes (Doyle et al. 2008). De même, on assiste à des réarrangements fonctionnels importants et à une modulation de l'expression des gènes (effets additifs chez les autopolyploïdes, effets non-additifs chez les allopolyploïdes, interactions épistatiques entre génomes homéologues) (Doyle et al. 2008). La polyploidie agit comme une force de mutation extrême permettant l'apparition rapide de nouvelles variations génétiques et phénotypiques. Par exemple, chez le coton, les espèces diploïdes de *G. arboreum* et *G. herbaceum* du groupe génomique A (espèces africaines), toutes comme les deux espèces tétraploïdes (AADD) américaines *G. barbadense* et *G. hirsutum* ont été domestiquées pour leurs capacités à produire de longues fibres de coton « filables » sur le tégument des graines. Les espèces américaines diploïdes du génome D produisent de fibres non filables (courtes) (Figure 1-5). Des études de cartographie génétique ont pourtant localisé la plupart des gènes impliqués dans la production de fibres de qualité au niveau du génome D, montrant ainsi l'importance des interactions épistatiques entre les deux génomes dans l'expression du principal trait de domestication (Jiang et al. 1998 cité dans Gepts 2004). De même chez la canne à sucre, il est notable que les deux ancêtres sauvages, *S. robustum* et *S. spontaneum* ont un contenu en sucre réduit, contrairement aux hybrides polyploïdes qui ont un contenu en sucre plus important (Grivet et al. 2004). Ces interactions géniques sont à l'origine d'innovations phénotypiques rapides servant de matériel de base aux cultivateurs pour la domestication mais aussi pour la diversification des formes cultivées. Il est important de noter que l'état polyploïde génère une intensification des mutations et des possibilités de variations, tout en permettant de supporter ou « tamponner » les potentiels effets négatifs de ces réarrangements de par la redondance génétique que cet état leur confère.

Aussi, la polyploïdisation permet en général d'élargir les potentialités écologiques des lignées polyploïdes par rapport à leurs progéniteurs diploïdes. Chez les allopolyploïdes, on réunit au sein d'un même génome des plantes pouvant présenter des adaptations écologiques différentes, élargissant ainsi le potentiel adaptatif du polyploïde. C'est clairement ce qui s'est passé pour le blé hexaploïde. Ce polyploïde combine le génome D de *A. tauschii* avec le génome AB du blé dur tétraploïde et présente ainsi une adaptation à une large gamme de conditions de photopériode et de vernalisation. Il a en outre une meilleure tolérance aux variations abiotiques du milieu (salinité, pH, température) et une meilleure résistance à de nombreux agents pathogènes (Dobscovski et al. 2007). Une revue récente suggère que l'autopolyploïdie pourrait être avantageuse dans des milieux « perturbés », là où les populations sont fragmentées et subissent de forts effets de dérive génétique et de la dépression de consanguinité (Parisod et al. 2010). Les variations climatiques sont souvent responsables de tels effets. La mise en culture lors de la domestication qui isole une fraction seulement de la population sauvage de départ induit des effets similaires, en particulier pour les plantes à propagation végétative et souvent allogames (Mc Key et al. 2010). L'état polyploïde et la ségrégation polysomique permettraient dans ces populations autopolyploïdes cultivées de limiter les effets de la dépression de consanguinité en maintenant une taille efficace de population plus importante.

### Figure 1-5 : Histoire évolutive du coton.

Deux espèces sauvages de coton diploïdes existent, l'une présente en Amérique tropicale et l'autre en Afrique. L'hybridation entre ces deux formes a eu lieu sur le continent américain il y a 1 à 2 millions d'années et a donné naissance à plusieurs espèces tétraploïdes. Les formes tétraploïdes *G. hirsutum* et *G. barbadense* ont alors été domestiquées indépendamment au Mexique et au Pérou, respectivement.



Enfin, la polyploïdisation est souvent associée à une baisse de la fertilité (Comai et al. 2005). Cela aurait pu permettre de limiter les flux de gènes entre les formes sauvages et cultivées, isolément reproducteur facilitant la transition du sauvage au cultivé et la fixation du syndrome de domestication. Dans certains cas, la baisse de fertilité a pu avoir des conséquences directes sur le syndrome de domestication. Des plantes qui investissent moins dans la reproduction sexuée sont susceptibles d'allouer plus de ressources à d'autres fonctions comme la mise en réserve au niveau du fruit ou d'organes souterrains (racines et tubercules) (voir partie II) (Mc Key et al. 2010). Ainsi, la domestication aurait favorisé des formes polyploïdes peu fertiles mais plus productives.

Pour bon nombre de plantes cultivées, les relations temporelles entre polyploïdisation et domestication demeurent incertaines (Hughes et al. 2007). Pour certaines plantes, il est clairement établi que la polyploïdisation est un événement ancien pré-datant la domestication (ex. du coton) (Wendel et al. 1995). Pour d'autres espèces, la polyploïdisation est directement liée aux perturbations exercées par les activités humaines de mise en culture et de domestication. En déplaçant des plantes (intentionnellement ou fortuitement), l'homme a mis en sympatrie des formes différenciées (espèces, sous-espèces ou populations) favorisant leur hybridation, parfois suivie d'événements de polyploïdisation. Hughes et al. (2007) ont récemment démontré l'importance de ce phénomène pour l'origine des formes domestiquées du genre *Leucaena* (arbres et arbustes d'Amérique centrale) et propose que ce même scénario soit aussi à l'origine des formes domestiquées des genres *Opuntia* et *Agave*. Ces hybridations ont pu avoir lieu assez tôt durant le processus de domestication ou seulement plus tard lors de la diffusion de la forme domestiquée.

L'homme ne se contente pas de domestiquer des plantes, mais il va progressivement les diffuser, les emmenant avec lui comme compagnon de traverse dans sa conquête des espaces, ou plus simplement les échangeant autour de lui au gré de réseaux d'interactions complexes. Ces mouvements qui se réalisent à des échelles spatiales et temporelles, variables selon l'espèce considérée, vont ainsi poursuivre le façonnement de la biodiversité cultivée.

## II. DIFFUSER DES PLANTES : LA DIVERSIFICATION DES PLANTES DOMESTIQUEES HORS DE LEUR FOYER D'ORIGINE

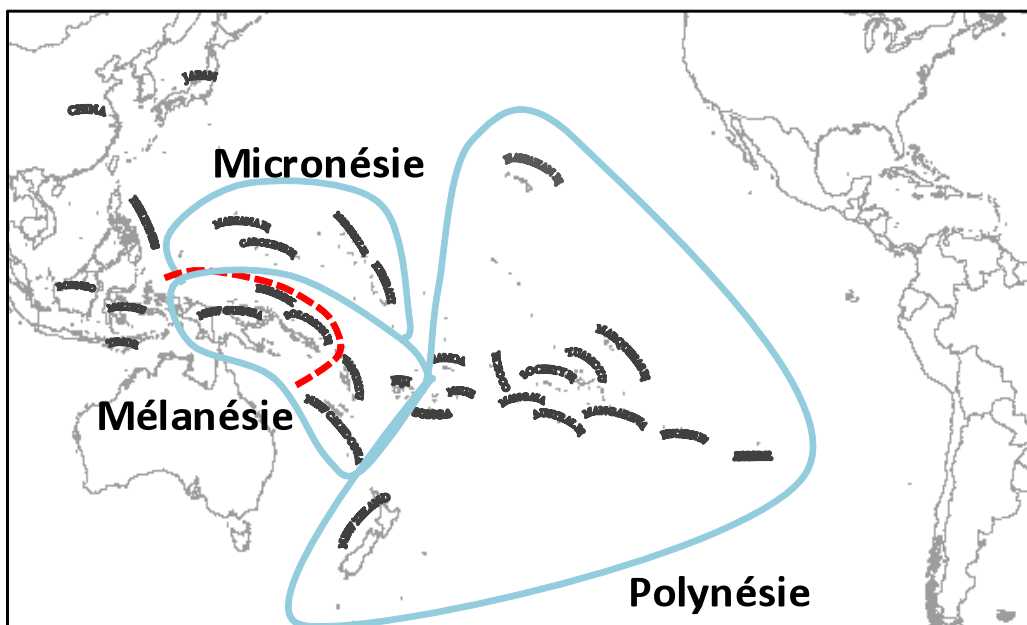
### II.1. La diffusion : un aperçu

#### II.1.1. Modes de diffusion

Selon l'hypothèse développée par Diamond et Bellwood (2003), l'adoption de l'agriculture dans certaines régions a conduit à une croissance démographique rapide des populations productrices de nourriture, les contraignant alors à migrer hors de leurs foyers de domestication originaux. L'agriculture aurait aussi permis le développement de sociétés hiérarchisées capables de se tourner vers la conquête de nouveaux espaces. D'importants mouvements migratoires auraient ainsi eu lieu durant l'Holocène. Ces vagues de colonisation auraient progressivement incorporé puis remplacé la plupart des chasseurs-cueilleurs, diffusant conjointement culture, langue, gènes, et plantes et animaux domestiques de la société colonisatrice. Ce modèle de diffusion est dénommé diffusion démique. Si cette hypothèse séduisante semble être appropriée dans certains cas, comme celui de l'expansion des populations d'agriculteurs de langue Bantou d'Afrique de l'Ouest en Afrique du Sud il y a environ 4 000 ans (Ehret 1998 cité dans Donohue et Denham 2010) ou encore l'expansion des Indo-européens en Eurasie (Ammermann et Cavalli Sforza 1984), les données génétiques, linguistiques et archéologiques ne concordent pas toujours et suggèrent pour d'autres cas des scénarios plus complexes (Donohue et Denham 2010).

### Encadré 1: L'espace Océanien

Traditionnellement, l'Océanie comprend l'ensemble des îles du Pacifique et exclue donc l'Australie ainsi que les îles d'Asie du Sud-est. On doit au navigateur français Dumont d'Urville, en 1930, le découpage de ce vaste territoire en trois zones : la Polynésie (« plusieurs îles »), la Mélanésie (« îles noires »), et la Micronésie (« petites îles »), découpage qui s'imposera jusqu'à très récemment et qui a longtemps structuré les recherches Océaniques, matérialisant des frontières là où il n'y en avait pas nécessairement. Dumont D'Urville élabore sa division tripartite à partir de considérations morales et physiques largement influencées par les visions européennes raciales et sexistes de l'époque, même s'il s'intéressa aussi aux langues et à l'organisation sociale et politique de ces populations. Ce découpage est remis en cause par les anthropologues et historiens dans la mesure où les groupes définis ne correspondent pas toujours à des entités culturelles et historiques identifiables. Il est aujourd'hui clair que la Mélanésie comme la Micronésie sont des ensembles hétérogènes, la Mélanésie étant même une des régions du monde hébergeant la plus grande diversité culturelle et linguistique, comparée à la taille du territoire. Seule la Polynésie, du fait de ses traditions maritimes au long cours, son unité linguistique et sa relative homogénéité sociologique et culturelle reste pertinente. Un autre découpage a alors été proposé, découpage qui reflète les dernières avancées concernant l'histoire du peuplement de l'Océanie. On parle ainsi d'Océanie proche pour délimiter la région qui englobe la Nouvelle Guinée, l'archipel Bismark et les îles Salomon (jusqu'à San Cristobal) (délimité par la ligne rouge sur la carte) et d'Océanie lointaine pour toutes les autres îles du Pacifique. L'Océanie proche a été peuplée dès le Pléistocène supérieure (il y a environ 40 000 ans) alors que le peuplement de l'Océanie lointaine ne commence qu'au milieu de l'Holocène, il y a environ 3 500 ans. Cette nouvelle division pourrait peut-être permettre de repenser les liens entre les îles d'Océanie proche et celles d'Asie du Sud-est. En effet, ces dernières sont la plupart du temps exclues de l'Océanie alors qu'il est maintenant clair qu'il y a eu de nombreuses interactions entre l'Océanie proche et les îles d'Asie du Sud-est avant même l'arrivée des Austronésiens.



### **Encadré 1 suite : Les Environnements du Pacifique**

L'Océan Pacifique est celui de la démesure : le plus large, il s'étend sur plus de 179 millions de km<sup>2</sup>, soit un tiers de la surface totale du globe; le plus profond, avec ses fosses océaniques vertigineuses qui peuvent atteindre 11000 km de profondeur (fosses des Mariannes) ; le plus ancien, âgé de plus de 190 millions d'année ; le seul reposant essentiellement sur un plaque tectonique, la plaque Pacifique. Cette immense étendue d'eau est parsemée de quelques 25 000 îles aux environnements et conditions écologiques très variées dont certaines seulement vont être peuplées par les hommes.

Certaines îles (dites continentales) sont des morceaux de croûtes continentales émergées, ou encore des arcs volcaniques à proximité d'une frontière entre deux plaques tectoniques convergentes. Dans le Pacifique, elles sont souvent de grande taille et se situent à l'Ouest de l'Océan (Nouvelle Guinée, archipel Bismarck, îles Salomon, Vanouatou, et Nouvelle Calédonie notamment), à relative proximité des continents asiatiques et australiens. Les autres îles, situées au milieu de la plaque Pacifique et donc souvent très éloignées des côtes, ont été créées par du volcanisme de point chaud. Lorsque ces îles sont récentes, elles forment des îles hautes, montagneuses (Tahiti, Rarotonga ou encore Hawaii). Lorsqu'elles vieillissent, s'érodent et s'affaissent, ces îles au faible relief dites atolls sont constituées de coraux affleurant à la surface de la mer. Ce type d'île constitue souvent un milieu aride et hostile à la vie végétale et humaine.

La plupart des îles du Pacifique connaissent des climats tropicaux ou subtropicaux où se distinguent une saison sèche et une saison humide, exceptée la Nouvelle Zélande où règne un climat tempéré. Localement, les hauts reliefs créent des microclimats montagnards plus froids. Les vents dominants (alizés) ainsi que les courants marins viennent majoritairement d'Est. Au niveau des îles hautes, ces vents créent des climats très contrastés entre les côtés « au vent » soumises à de fortes précipitation et celles « sous le vent » largement plus sèches. Les îles du Pacifique sont soumises de manière chronique ou ponctuelle à de nombreuses catastrophes naturelles, ouragans, tremblements de terre, tsunamis et éruptions volcaniques. Aussi, le phénomène climatique *el Nino* vient cycliquement bouleverser les climats Océaniens, provoquant notamment d'importantes périodes de sécheresse dans l'ouest de l'Océan Pacifique.

Ces caractéristiques géomorphologiques et climatiques structurent localement les patrons biogéographiques des îles du Pacifique. De façon globale, on observe un gradient de diversité d'Ouest en Est. L'Ouest du Pacifique possède une flore et une faune riche représentés par de nombreuses familles et genres des continents limitrophes, qui recèle potentiellement d'espèces utilisables par l'homme. A l'inverse, la diversité biologique des petits archipels polynésiens à l'Est est réellement appauvrie. A ce gradient de diversité s'ajoute aussi un gradient de vulnérabilité lié notamment à l'isolement croissant des îles à l'Est du Pacifique.

Les hommes qui peupleront le Pacifique d'Ouest en Est devront donc au fil des îles s'adapter à ces contraintes environnementales « grandissantes ».



**Figure 1-6: Les principales plantes du complexe horticole océanien. (© V. Lebot)**

**a) Plantes à racines et tubercules**

**La grande igname (*D. alata*)**



**Tubercules de grandes ignames**



**Bulbes de l'igname *D. bulbifera***



**Tubercules de l'igname *D. nummularia***



**Tarodières de *C. esculenta***



**Cormes de *C. esculenta***



**Cormes d'*A. macrorrhiza***



Ainsi, d'autres modèles défendent plutôt l'idée d'une diffusion dite culturelle (Zvelebil et Zvelebil 1988), de proche en proche, par influence, au sein de réseaux d'interactions régionales, mais n'ayant pas impliqué d'importants mouvements de populations. Contrairement à la diffusion démique où les agriculteurs emmèneraient dans leurs conquêtes des espaces un ensemble de plantes et animaux domestiques, un « paquet agricole », la diffusion culturelle donne place à des échanges multidirectionnels. Ainsi les différentes plantes cultivées sont échangées une à une et connaissent des histoires évolutives individuelles (Donohue et Denham 2010). Ces deux modes de diffusion ne sont bien entendu pas mutuellement exclusifs, et leur combinaison est tout à fait envisageable dans certaines régions.

L'époque post-colombienne a ensuite largement étendu et intensifié cette diffusion, aboutissant à une véritable mondialisation des échanges et une recomposition des systèmes agricoles qui intégreront tant bien que mal ces plantes venues d'ailleurs.

### *II.1.2. Echelles spatio-temporelles*

Diamond (2003) propose que la vitesse de diffusion de l'agriculture « primitive » ait été largement conditionnée par l'orientation des continents. Ainsi sur le continent eurasiatique, la diffusion aurait pu être accomplie assez rapidement selon un axe latitudinal est-ouest, où les plantes ne sont pas soumises à de fortes variations des conditions climatiques et écologiques. En revanche, l'orientation longitudinale nord-sud des continents américains (ou africain) aurait ralenti le processus de diffusion, les plantes devant s'adapter en chemin à des conditions climatiques très variables. Il est certain que la latitude n'est pas le seul déterminant et les barrières topographiques (montagnes, océans) et écologiques sont autant d'obstacles locaux à la diffusion. Les caractéristiques intrinsèques de la plante (large adaptabilité ou exigences écologiques restrictives, par exemple), tout comme les préférences ethniques, ont aussi conditionné les patrons de diffusion des espèces cultivées. On attribue souvent aux voyageurs européens (dès le 15<sup>ème</sup> siècle) l'intensification de la diffusion des plantes cultivées à grande échelle. Pourtant, des études récentes alliant génétique et archéologie ont pu mettre en évidence l'existence d'anciens réseaux d'échanges entre l'Afrique, la péninsule arabique et l'Asie du Sud, attestant déjà d'échanges sur de grandes distances (Zeder et al. 2006, Fuller et al. 2012). De même, la distribution de certaines plantes laisse penser qu'il a dû y avoir des voyages trans-océaniques anciens, impliquant notamment des échanges préhistoriques de plantes cultivées entre l'ancien et le nouveau monde (Yen 1974, Erickson et al. 2005, Green 2005, Clarke et al. 2006, Baudoin et Lebrun 2009).

### *II.1.3. Conséquences génétiques et évolutives de la diffusion des plantes domestiquées*

La diffusion représente en général une succession de goulots d'étranglement qui participent à réduire la diversité génétique des plantes cultivées dans les zones d'introduction par rapport à leurs centres d'origine. Des introductions multiples ont dans certains cas atténué la sévérité des premiers effets de fondation. Parfois, la diffusion achève la domestication en permettant d'isoler véritablement la forme cultivée de ses ancêtres sauvages. Elle permet aussi de mettre en contact les formes cultivées avec de nouvelles formes sauvages, tout comme de réunir des formes cultivées domestiquées indépendamment, qui vont pouvoir s'hybrider et ainsi ré-enrichir le compartiment cultivé. Dans certains cas, ces hybridations - plutôt que de régénérer de la diversité -, peuvent conduire à la production et à la fixation par la sélection de formes stériles de la plante en question, limitant par la suite son potentiel évolutif. Les effets de la diffusion sur la diversité génétique d'une plante cultivée peuvent donc être divers.

La diffusion permet aussi d'exposer les plantes cultivées à de nouvelles conditions environnementales, climatiques, écologiques, mais aussi biotiques, qui vont exercer autant de nouvelles pressions de sélection sur les espèces introduites. Les véritables garants de cette diversification sont bien sûr les paysans eux-mêmes. Les pratiques paysannes tiennent une place centrale dans la structuration et la dynamique de la diversité des plantes cultivées au cours de leur diffusion. Elles se combinent aux caractéristiques biologiques des plantes (systèmes de reproduction notamment) et aux pressions environnementales pour influencer la plupart des forces évolutives (dérive, sélection et migration). Des études menées dans différents agro-systèmes traditionnels mettent



b) Arbres, arbustes et arbrisseaux

**Feuille et fruit de l'arbre à pain**



**Choux des îles**



**Tige et arbuste de kawa**



**Pandanus**



**Nangailier**



en évidence des pratiques paysannes qui valorisent la diversité, ainsi que l'intérêt des paysans à maintenir au cours du temps cette diversité agricole (ex. Boster et al. 1984, Brush et al. 1995, Salik et al. 1997, Elias et al. 2000, Perales et al. 2005, Scarcelli et al. 2006, Caillon et al. 2006, Barnaud et al. 2007, Sardos et al. 2008). Dans ces agro-systèmes, cette diversité assure en premier lieu aux agriculteurs la durabilité et la résilience de leurs systèmes vivriers face aux aléas environnementaux et aux pressions imposées par les agents pathogènes (Altieri 1999). Cette diversité biologique est aussi valorisée en tant que telle comme expression d'une nécessité culturelle (Caillon et Degeorges 2007). Aussi, est-il important de noter que la diversité phénotypique - celle que sélectionne les agriculteurs, n'est pas toujours le reflet de la diversité génétique sous-jacente - celle qui garantit pourtant le potentiel évolutif de la plante (ex. Meilleur 1998, Zerega et al. 2004).

Pour une plante cultivée donnée, la diffusion crée donc des patrons hétérogènes de distribution de la diversité, diversité phénotypique comme génétique: certaines régions sont clairement appauvries par rapport à la zone d'origine, alors que d'autres apparaissent comme de véritables points chauds de diversité, appelés couramment centres de diversité secondaires (Pickersgill 1998). Alors que le fonctionnement des centres d'origine de bon nombre de plantes est relativement bien caractérisé, celui des centres de diversité secondaire l'est beaucoup moins. Sont-ils des réservoirs passifs de la diversité existante dans d'autres régions, ou des lieux de diversification active qui accumulent de nouveaux caractères ? Quels sont les processus à l'origine de cette concentration de diversité ? Plusieurs cas d'étude, comme le haricot commun en Afrique de l'Est, le manioc en Afrique de l'Ouest, ou encore l'orge en Ethiopie ont permis de lever le voile sur certains des processus qui favorisent l'émergence de ces centres de diversité secondaires (Pickersgill 1998). Ainsi, l'incroyable diversité de haricots présents en Afrique résulterait à la fois de la concentration de la diversité présente dans les deux zones d'origine américaines (mésosud-américaines) et de pratiques de culture poly-variétales, plus sophistiquées que celles pratiquées dans les zones d'origine.

L'histoire des migrations des hommes et de la diffusion des plantes en Océanie (encadré 1) est un exemple particulièrement riche pour appréhender les modalités de cette diffusion, ainsi que ses conséquences évolutives sur les patrons de diversité des plantes cultivées. Les Océaniens pratiquent une agriculture vivrière particulière, dénommée horticulture<sup>1</sup>. Ses systèmes horticoles sont centrés sur la culture et l'association de plantes à multiplication végétative ou végéculture (Barrau 1962), et utilisent principalement des espèces comme les ignames, le taro, d'autres aracées (*Alocasia macrorrhiza* et *Cyrtosperma chamissonis*), mais aussi la patate douce, le manioc, le macabo (*Xanthosoma sagittifolium*) ou d'autres plantes, comme le chou des îles (*Abelmoschus manihot*), le kava, la banane ou la canne à sucre (*Saccharum* sp.) (Walter et Lebot 2007) (Figure 1-6). Le deuxième pivot de l'agriculture océanienne est l'arboriculture, où des espèces comme l'arbre à pain, le nangailler (*Canarium* sp.), les pandanus (*Pandanus* spp.), les cocotiers (*Coco nucifera*), ou encore les sagoutiers sont cultivés ou simplement entretenus dans les forêts alentours pour leurs graines et leurs fruits (Walter et Lebot 2007) (Figure 1-6). A la diversité des environnements océaniens (encadré 1) correspondent des associations vivrières spécifiques: aux forêts humides tropicales des systèmes vivriers basés sur la culture du taro; aux zones de forêts plus sèches et de savanes des systèmes à base d'ignames; et aux régions montagnardes des hautes terres de Nouvelle-Guinée une culture vivrière à base de patates douces (Barrau 1962). Ces systèmes vivriers océaniens sont en général caractérisés par une forte diversité spécifique, ainsi que par une importante diversité intra-spécifique ou diversité variétale (Lebot 1992 et 2002). Cette double diversité assure depuis l'installation des premiers hommes la résilience du système agricole dans ces îles vulnérables, isolées, et soumises de manière chronique ou ponctuelle à de fortes variations environnementales et de nombreuses catastrophes naturelles (Lebot 1992). La plupart des espèces cultivées en Océanie sont propagées par voie clonale par les agriculteurs. D'où provient donc cette impressionnante biodiversité cultivée, et qu'elle en est la

<sup>1</sup> Pour J. Barrau (1967), la différence entre l'**agriculture** et l'**horticulture** réside à la fois dans la dimension de l'unité agraire, l'enclos de l'*hortus* et l'étendu de l'*ager*, et dans les pratiques culturelles mises en œuvre dans les deux systèmes. Les plantes agricoles sont semées à la volée, récoltées à la faux, alors que les végétaux horticoles sont plantés uns à uns, repiqués, bouturés puis entourés de soins individuels. La **végéculture** se définit comme l'agriculture ou l'horticulture des plantes à multiplication végétative.



c) « *Les herbes* »

**Canne à sucre**



***Cordyline fruticosa***



**Bananier**



base génétique réelle? Elle est le fruit d'une longue histoire de migrations des hommes d'îles en îles, de plantes sur des pirogues, d'échanges aussi à travers l'océan sur des voiliers européens, une histoire que nous allons nous attacher à retracer ici dans ses grands traits.

## **II.2. La diffusion des hommes et des plantes en Océanie**

### ***II.2.1. Une origine des plantes océaniques plutôt sur Sahul que Sunda***

L'expansion rapide des Austronésiens, initiée il y a environ 4500 ans depuis Taiwan jusque dans les îles du Sud-est asiatique puis à l'est dans le Pacifique est souvent citée comme l'archétype d'une diffusion démique (Diamond et Bellwood 2003, Bellwood 2011). Dans la lignée d'un tel modèle, on a longtemps considéré que les plantes et animaux domestiques d'Océanie avaient une origine asiatique (Sauer 1952), tout comme les hommes qui les transportaient. Récemment, la compilation de données archéologiques, linguistiques et moléculaires a fait voler en éclats ce paradigme, démontrant que la Nouvelle-Guinée était un des plus anciens berceaux de l'agriculture (Yen 1991, Lebot 1999, Pickersgill 2004, Allaby 2007, Denham 2010) et que le peuplement de l'Océanie était un processus graduel ayant impliqué de nombreuses interactions avec les populations préexistantes (Kirch 2000 et 2010, Donohue et Denham 2010).

### ***II.2.2. Les premiers migrants du Pléistocène en "Océanie proche"***

Une première vague d'hommes modernes (*Homo sapiens sapiens*) a gagné l'"Océanie proche" dès le Pléistocène supérieur. A cette époque, le niveau des mers était plus bas et la géographie des continents différente (Figure 1-7). A l'Ouest, la Nouvelle-Guinée, l'Australie et la Tasmanie formaient un unique continent, dit Sahul, séparé de l'archipel des Salomon par une frange d'eau. A l'est, la Péninsule Malaise était reliée à l'arc des îles indonésiennes, formant Sunda (la plaque continentale du Sud-est Asiatique). Entre ces deux continents se tenait le Wallacea (ainsi nommé parce que positionné sur la « ligne de Wallace »), un espace insulaire constamment immergé, faisant office de barrière à la dispersion des plantes, des animaux, mais aussi des hommes (Figure 1-7). Plus de 150 sites archéologiques datant du Pléistocène supérieur (Kirch 2000 et 2010) ont été découverts sur Sahul ainsi que sur les îles d'"Océanie proche" (actuels archipel Bismarck et îles Salomon). L'exploration de ces sites montre que des migrants venant d'Asie ont traversé *Wallacea* il y a entre 60 000 et 40 000 ans, probablement sur des radeaux de bambous, puis ont gagné rapidement Sahul ainsi que les îles d'Océanie proche dès 36 000 avt Jésus Christ.

### ***II.2.3. De la cueillette aux débuts de l'agriculture en Nouvelle-Guinée***

Ces premiers migrants du Pléistocène vivaient essentiellement de chasse, de pêche et de cueillette, comme tous les groupes humains de cette époque. Assez tôt (il y a 20 000 ans), ces populations ont commencé à intensifier leurs pratiques d'exploitation du milieu, à aménager et jardiner les forêts environnantes, pratiquant ainsi les premières formes d'arboriculture (Yen 1995 et 1996). Ils déplaçaient aussi des animaux, comme par exemple le cuscous (*Phalanger orientalis*) ou le bandicoot (*Epymichera calubu*) depuis la Nouvelle-Guinée jusqu'aux îles environnantes pour la chasse (Irwin 2006). « L'agriculture » à proprement parler a débuté en Nouvelle-Guinée assez tôt durant l'Holocène, de manière concomitante avec les autres foyers de domestication identifiés ailleurs dans le monde. C'est principalement l'exploration de divers sites archéologiques, notamment un site dénommé « Kuk » situé dans les zones marécageuses de la vallée de Waghi des hautes terres occidentales de Papouasie-Nouvelle-Guinée, qui a permis de révéler l'existence et l'évolution des pratiques agricoles dans cette région dès 10 000 avt J.C. (Golson 1980, Denham et al. 2004 et 2010, Golson 2007). Les premières formes d'agriculture semblent avoir été menées sur brûlis, et consistaient probablement en la mise en culture dans ces espaces dégagés de formes sauvages d'arbres, de bananiers (*Musa*



d) *Plantes introduites d'Amérique du Sud*



**Champ et tubercules  
de patate douce**



**Gourde**

*acuminata* ssp. *banksii*) et probablement de taros. Dès 5 000 ans avant JC, l'intensification agricole se matérialise par l'apparition d'une agriculture sur buttes. D'abondants restes archéologiques de bananiers (phytolithes de *Musa* sp.) et de taros (grains d'amidon de *Colocasia* sp.) sont identifiés. Puis vers 2 000 avant JC se développent des systèmes de drainage des marécages par le creusement de fossés. Dans les zones tropicales humides, les archéologues peinent souvent à identifier des restes botaniques de plantes à multiplication végétative qui se dégradent très rapidement, et lorsqu'ils en identifient (phytolithes, grains d'amidon ou parenchymes), ils peinent à savoir s'il s'agit de formes sauvages ou domestiquées (Denham 2007). Les données moléculaires ont donc été capitales pour appuyer l'idée d'un centre de domestication en Nouvelle-Guinée (Lebot 1999 et 2002, Allaby 2007). L'identification de deux groupes génétiques distincts de taro, l'un correspondant aux cultivars d'Asie du Sud-est, et l'autre à ceux du Pacifique (notamment la Nouvelle-Guinée), suggère fortement une domestication dans chacune de ces régions à partir de populations locales déjà différenciées génétiquement (Lebot et Aradhya 1999, Kreike et al. 2004, Lebot et al. 2004). De même, les données moléculaires montrent que la sous-espèce *Musa acuminata* ssp. *banksii* (forme diploïde), endémique de Nouvelle-Guinée et toujours exploitée aujourd'hui, a contribué aux génomes de la plupart des autres cultivars de bananiers dispersés dans le monde, suggérant que la domestication a été initiée en Nouvelle-Guinée (Lebot et al. 1994, De Langhe et De Marret 1999, Perrier et al. 2011). De même pour l'arbre à pain (*Artocarpus altilis*), les patrons de diversité obtenus dans le Pacifique suggèrent que les formes domestiquées de Mélanésie et Polynésie dérivent d'une espèce indigène de Nouvelle-Guinée, *A. camensis* (Zerega et al. 2004).

#### **II.2.4. Des réseaux d'interactions précoces avec les îles d'Asie du Sud-est**

Une étude récente couplant des données de génétique et de linguistique a permis de retracer pas à pas l'histoire évolutive de la banane (Perrier et al. 2011). La diversité des cultivars modernes de banane serait le résultat de l'hybridation de *Musa acuminata* ssp. *banksii*, la sous-espèce domestiquée en Nouvelle-Guinée, avec d'autres sous-espèces des îles d'Asie du Sud-est, ainsi qu'avec une autre espèce, *Musa balbisiana* originaire d'Asie continentale (Perrier et al. 2011). Ces transferts de végétaux d'est en ouest, réalisés nécessairement par des hommes, remonteraient à plus de 4 000 ans, pré-datant ainsi la deuxième vague de migrants, les Austronésiens, qui peupleront par la suite l'Océanie lointaine depuis Taiwan. Ce même scénario est envisagé pour la canne à sucre (*Saccharum* sp.) (Grivet et al. 2004). La première forme domestiquée à partir de *S. robustum* en Nouvelle-Guinée aurait été introduite en Asie du Sud-est et se serait hybridée localement avec l'espèce *S. spontaneum*, produisant de nouveaux cultivars, riches en sucre, regroupés sous le nom de *S. officinarum*. De même, une étude génétique récente propose que le porc, initialement domestiqué en Asie du Sud-est, ait été introduit précocement en Nouvelle-Guinée (Larson et al. 2007). Ces nombreux travaux de phylogéographie des espèces domestiquées mettent ainsi en évidence l'existence d'interactions anciennes (durant l'Holocène mais pré-datant l'arrivée des Austronésiens), et probablement nombreuses entre l'Océanie proche et les îles d'Asie du Sud-est. Ces interactions sont aussi suggérées par certaines données de génétique humaine et de linguistique (Donohue et Denham 2010).

Contredisant le paradigme initial diffusionniste en faveur de l'importation complète d'un paquet agricole en "Océanie proche" depuis un centre de domestication asiatique, les données actuelles révèlent donc une image plus complexe. En effet, certaines espèces ont été domestiquées en Nouvelle-Guinée puis dispersées à l'ouest, d'autres domestiquées à l'ouest puis diffusées vers l'est, et d'autres encore, domestiquées localement et indépendamment dans les deux régions et n'ayant pas fait l'objet d'échanges entre elles (Tableau 1-1). Ces événements précoces de diffusion de plantes déjà domestiquées ont mis en contact très tôt des populations isolées depuis longtemps de part et d'autres de Wallacea et génétiquement différenciées (Donohue et Denham 2010). Ces hybridations interspécifiques ont eu des conséquences évolutives majeures sur les espèces en question. Les graines de bananiers sont nombreuses et très dures, rendant le fruit peu comestible. Très tôt, en Nouvelle-Guinée, la sélection humaine a favorisé des formes sans graines parthénocarpiques. Puis, ces hybridations interspécifiques ont largement contribué à l'obtention et à la sélection de nouvelles formes stériles diploïdes et polyploïdes avec une plus grande vigueur et des fruits volumineux, remplaçant alors les diploïdes initiaux à petits fruits, hormis dans la zone d'origine en Nouvelle-



**Figure 1-7 : Les terres émergées Sunda et Sahul (en vert) il y a environ 50,000 ans lorsque les premiers migrants du pléistocène ont peuplé la Nouvelle Guinée depuis l'Asie. La ligne rouge représente la ligne de Wallace. Ces deux masses de terre étaient séparées par un espace insulaire constamment immergé, faisant office de barrière à la dispersion des plantes, des animaux, mais aussi des hommes. (D'après Bourke et Harwood 2009)**



Guinée. La plupart des cultivars qui vont diffuser ensuite seront ces formes stériles polyploïdes, pour lesquelles la seule source de diversification reste la mutation somatique et les modifications épigénétiques (Lebot 2002, Perrier et al 2011).

## ***II.2.5. L'arrivée des austronésiens en Océanie proche et le peuplement de l'Océanie lointaine***

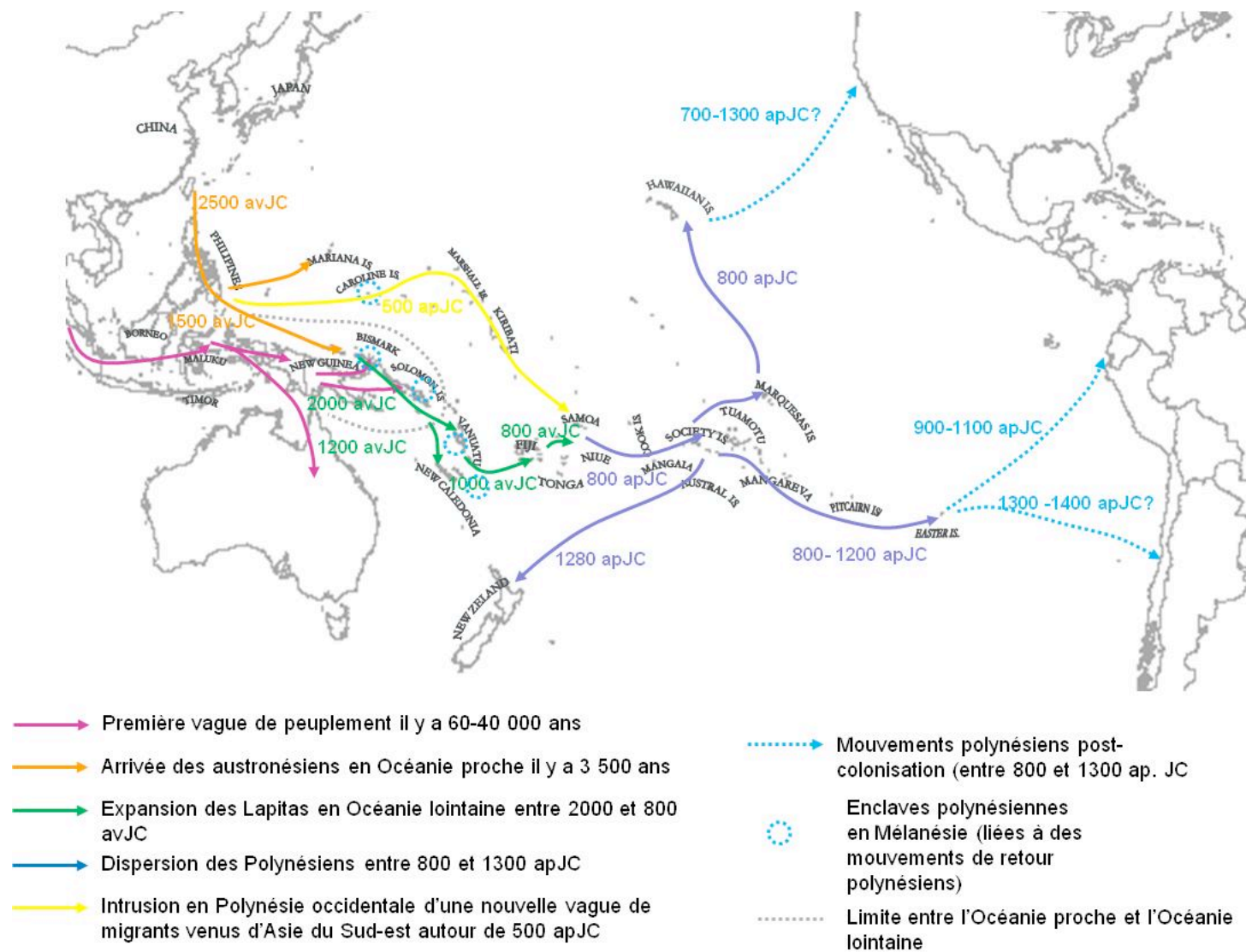
### *L'arrivée des austronésiens*

Les données archéologiques suggèrent qu'une seconde vague de migrants a colonisé l'"Océanie proche" il y a environ 4 000 ans, principalement sur les îles de l'archipel Bismarck, le nord des îles Salomon et dans une moindre mesure les côtes de Nouvelle-Guinée (Kirch 2000 et 2010) (Figure 1-8). Ces nouveaux arrivants sont caractérisés par leurs poteries de type « Lapita », absentes jusqu'ici des sites archéologiques et par leurs systèmes de navigation, les pirogues à balancier (« wanka »). Ces populations parlent des langues dites austronésiennes, aujourd'hui répandues dans toute l'Asie du Sud-est, dans certains sites côtiers de Nouvelle-Guinée et dans toute l'Océanie. Elles auraient quitté Taiwan il y a 5 000 ans, seraient passées par les Philippines et les autres îles du Sud-est asiatique avant d'atteindre les îles de "Océanie proche" il y a 3 500 ans et enfin de s'élancer vers "Océanie lointaine".

L'analyse comparée des données de linguistique et de génétique humaine montre qu'au cours du processus de diffusion, il y eu des échanges de gènes entre les migrants et les primo-résidents (issus des premières migrations du Pléistocène) (revue par Keiser 2010). En "Océanie proche", les données de génétique humaine obtenues à partir de l'étude des marqueurs mitochondriaux concordent souvent avec les appartenances linguistiques et montrent que les groupes d'affinité linguistique austronésienne ont majoritairement une constitution génétique héritée de celles des populations taïwanaises, constitution différente de celle des groupes « Papous » de l'intérieur des terres de Nouvelle Guinée. L'analyse de la section non recombinante du chromosome Y (marqueurs NRY) montre cependant des patrons régionaux plus contrastés où certains groupes d'affinité linguistique austronésienne ont pourtant une constitution génétique mixte, parfois même majoritairement héritée des groupes Papous (Kayser 2010). Ce même signal génétique contradictoire et lié au sexe se retrouve dans toute l'Océanie. En Polynésie notamment, 96% des lignées mitochondriales sont originaires d'Asie de l'Est, alors que 62 % des lignées NRY proviennent de Nouvelle-Guinée. L'analyse de marqueurs autosomaux confirme cette double ascendance de l'ensemble des populations austronésiennes d'Océanie à des degrés différents en "Océanie proche" et en "Océanie lointaine" (Wollstein et al. 2010). Ces patrons génétiques sont interprétés comme la conséquence de « mariages asymétriques » entre migrants et primo-résidents, ayant eu lieu très probablement en "Océanie proche" (archipel de Bismarck), et ayant impliqué des hommes « locaux » et des femmes « migrantes », en relation avec les systèmes de résidence matrilocaux des sociétés proto-océaniques (Jordan et al. 2009)

Ces populations n'ont pas seulement échangé des gènes mais aussi leurs objets culturels. Les Austronésiens ont probablement apporté sur leurs pirogues leurs plantes et animaux domestiques issus du foyer de domestication chinois, notamment le riz et le porc. Quelques rares traces de riz ont pu être identifiées en Asie du Sud-est mais jamais en Océanie (Donohue et Denham 2010, Bellwood 2011). Il est probable que les Austronésiens aient rapidement abandonné la culture du riz dans ces régions équatoriales, en faveur de l'économie locale performante fondée sur l'exploitation des arbres et des tubéreuses (Bellwood 2011). Cette vague migratoire aurait tout de même introduit de nouvelles plantes utiles en Océanie comme le jamalaquier (*Syzygium malaccense*), le châtaignier de Tahiti (*Inocarpus fagifer*) ou encore le mûrier à papier (*Broussonetia papyrifera*), participant ainsi à la constitution d'un paquet horticole d'une centaine de plantes, qui sera ensuite diffusé en "Océanie lointaine" (Allaby 2007, Walter et Lebot 2007) (Figure 1-6). De ces interactions faites d'intrusions, d'intégrations mais aussi d'innovations, émerge le groupe culturel « Lapita » (Green 2003), nommé d'après les poteries du même nom (trouvées sur le site archéologique éponyme de Lapita près de Koné, dans le nord de la Nouvelle-Calédonie).

Figure 1-8 : Vagues de peuplement de l'Océanie et mouvements post-colonisation des Polynésiens.



Les données archéologiques attestent que le peuplement de l'"Océanie lointaine" s'est réalisé en deux étapes (Figure 1-8). Des sites archéologiques caractéristiques de la culture Lapita ont été retrouvés aux îles Reef (groupe des Santa Cruz, datés par le radiocarbone à 3 200 ans), au Vanouatou (daté à 3 100 ans), en Nouvelle-Calédonie (il y a 2 900 ans), aux îles Fidji (il ya 3 100 ans), et aux îles Tonga et Samoa (il y a 2 850 ans), attestant d'une extension rapide jusqu'en Polynésie occidentale (Kirch 2000 et 2010). Dans cette région, une pause d'au moins un millier d'années aurait été marquée, temps au cours duquel émergera la culture Polynésienne. Le groupe polynésien forme aujourd'hui une unité génétique, linguistique et culturelle, attestant d'une origine unique en Polynésie occidentale, et d'un processus de colonisation rapide. Il est possible que d'autres migrations en provenance d'Asie du Sud-est, migrations passées par les îles Caroline, les Kiribati puis les Tuvalu, soient aussi arrivées en Polynésie occidentale et aient participé à l'émergence de la culture polynésienne (Addisson et Matisso Smith 2010) (Figure 1-8). La chronologie exacte de cette colonisation est toujours en discussion. Suivant des nouvelles datations récentes, les polynésiens auraient quitté la région des Tonga/Samoa puis peuplé la chaîne d'îles de Polynésie orientale, des Australes à l'Île de Pâques, assez rapidement il y a 800 à 1000 ans. Ils auraient alors atteint Hawaï il y a environ 1 200 ans, en passant par les îles Marquises (Kirch 2010). Enfin, la colonisation de la Nouvelle-Zélande aurait eu lieu plus tardivement, il y a seulement 800 ans (Wilmschurt et al. 2008).

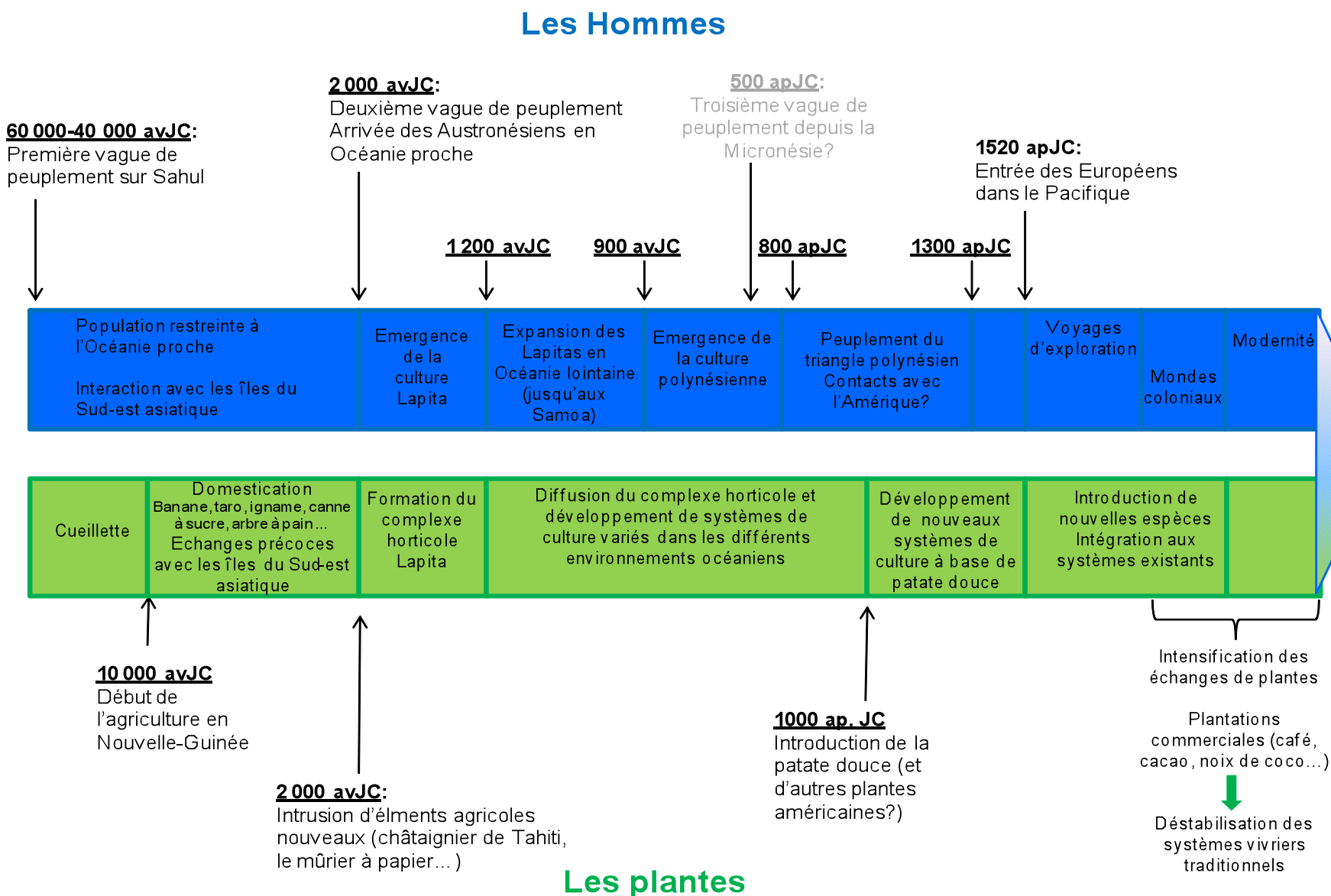
Peupler l'"Océanie lointaine" impose de traverser des étendues d'eau de plus en plus grandes, de plusieurs centaines à plusieurs milliers de kilomètres, alors qu'en "Océanie proche" il est souvent possible de naviguer à vue. Il faut aussi naviguer souvent contre les vents et courants dominants venant d'est. Les Lapitas possédaient déjà des systèmes de navigations avancés, modèles perfectionnés des premières pirogues à balanciers des Austronésiens. Les Polynésiens utiliseront plutôt de grands catamarans à voile (pirogues simples accolées, ou solidement liées), plus stables pour leurs longues navigations hauturières (Irwin 2006).

Peupler l'Océanie lointaine implique aussi de réussir à s'implanter sur des îles, parfois petites, de plus en plus vulnérables et éloignées du continent, qui possèdent une flore et une faune qui s'appauvrit progressivement d'ouest en est en genres et familles utiles (seulement 31 espèces de plantes à fleurs identifiées sur l'Île de Pâques), avec une rupture importante notamment à partir des îles du centre du Pacifique. Il est pourtant remarquable que sur l'ensemble des îles du Pacifique Sud, de la Mélanésie à la Polynésie, les principales espèces cultivées sont identiques (Lebot, 1992, 2002). Les Lapitas possédaient des systèmes horticoles développés et exploitaient probablement plusieurs centaines d'espèces végétales différentes (Kirch 2000). D'îles en îles, ils emmenaient avec eux les éléments biologiques et culturels qui leur permettent de reconstruire l'environnement qu'ils viennent de quitter. Ils transportaient leurs paysages. Les polynésiens poursuivront à leur tour cette dispersion des plantes et animaux domestiques depuis la Polynésie occidentale. L'introduction de ces différentes plantes a finalement donné naissance à des agro-écosystèmes variés le long des îles du Pacifique, où certaines plantes ont pris localement plus d'importance que d'autres, en fonction des facteurs environnementaux (climatiques et édaphiques), mais aussi biologiques et culturels (Figure 1-6 et 1-9). Ce même phénomène se retrouve parfois à l'échelle d'une île, notamment pour les îles hautes, en relation avec l'existence d'un climat contrasté, où la zone exposée au vent est humide et plutôt favorable à l'implantation d'un agro-écosystème basé sur l'exploitation du taro, tandis que la zone sous le vent est plus sèche et propice à la culture de l'igname (Barrau 1965).

#### ***II.2.6. Des contacts préhistoriques entre les Polynésiens et les Américains?***

La possibilité de contacts préhistoriques entre Polynésiens et Américains a ressurgi plusieurs fois dans les débats sur l'origine des hommes du Pacifique, et ce depuis plus de deux cents ans. Thor Heyerdahl, anthropologue, archéologue et navigateur norvégien, monte en 1947 l'expédition du Kon-Tiki, comme une tentative de rallier les îles polynésiennes sur un radeau depuis les côtes d'Amérique du Sud, pour démontrer que le peuplement de la Polynésie a pu avoir lieu depuis l'Amérique. Aucune preuve à ce jour n'est venue étayer la théorie de Heyerdahl. Cependant, un faisceau de preuves suggère qu'il y a

Figure 1-9 : L'histoire croisée des Hommes et des plantes en Océanie.



bien eu des connections anciennes entre ces populations, mais davantage attribuées aux voyages de Polynésiens sur les côtes américaines (revue par Jones et al. 2011).

Des preuves culturelles sont évoquées très tôt pour justifier de tels contacts. Parmi elles, on peut citer notamment l'existence de pirogues faites de planches cousues et d'outils de pêche de style polynésien, retrouvés dans les sites archéologiques des îles Channel dans le Sud de la Californie (Jones 2011), ou encore la ressemblance de certains outils des zones Mapuche du centre-sud du Chili et de Polynésie (Ramirez-Aliaga 2011). La génétique humaine a aussi été convoquée pour la recherche de telles preuves mais elle s'est souvent avérée peu résolutive. Les populations polynésiennes et celles d'Amérique ont des signatures génétiques bien distinctes. Certains traceurs américains du chromosome Y ont aussi été trouvés en faible fréquence en Polynésie, notamment sur l'île de Rapa (groupe des îles Australes), mais des données historiques y révèlent en 1863 la présence d'un bateau péruvien impliqué dans la traite d'esclaves polynésiens, bateau qui auraient pu laisser des descendants (Hurles et al. 2003). Une publication récente identifie des haplotypes HLA (Antigènes des Leucocytes Humains) spécifiques des Amérindiens sur l'Île de Pâques. En combinant ces données génétiques aux données généalogiques, les auteurs concluent que ces haplotypes HLA ont dû être introduits très tôt, probablement avant l'arrivée des Européens, alimentant ainsi le débat de nouvelles controverses (Lie et al. 2007, Thorsby 2012).

Ces contacts préhistoriques n'ont pas nécessairement impliqué d'échanges de gènes entre populations humaines et sont donc difficilement évaluables par la génétique humaine. En revanche, de tels contacts peuvent être pistés au travers de l'étude de la distribution et de la phylogéographie des espèces domestiquées qui auraient été transportées et échangées durant ces voyages (Storey et al. 2011). En effet, c'est principalement la présence ancienne d'espèces domestiquées américaines dans les systèmes vivriers océaniques, tout comme celles d'espèces asiatiques en Polynésie, qui a donné du poids à cette théorie.

#### *Le poulet asiatique en Amérique*

Il était admis que le poulet (*Gallus gallus*) domestiqué en Asie aurait ensuite été introduit en Amérique par les Européens. Cependant, la découverte d'une race très particulière dite « Araucana », trouvée au Chili, suggérait déjà d'autres sources d'introduction potentielles. Récemment, des ossements de poulet ont été retrouvés dans un site archéologique côtier du sud du Chili (site El Arenal) et datés précisément à 1304-1424 ans après J.C. (Storey et al. 2007). L'analyse de l'ADN de ces échantillons a montré qu'ils partageaient le même haplotype mitochondrial que d'autres restes anciens retrouvés à Tonga et au Samoa suggérant une introduction précoce en Amérique, en provenance de Polynésie (Storey et al. 2007). En 2008, une seconde étude met en cause les datations des ossements du site et montre que cet haplotype est largement partagé par la plupart des races commerciales actuelles, et aurait pu être introduit plus tardivement (Gongora et al. 2008). Depuis, d'autres ossements du même site ont été identifiés et datés, confirmant la présence préhistorique du poulet en Amérique du Sud (Storey et al. 2008).

#### *Des plantes cultivées américaines en Océanie*

La gourde (*Lagenaria siceraria*) aurait pu être amenée d'Amérique par les Polynésiens (Clarke et al. 2006). Les gourdes actuelles correspondent à deux groupes morphologiques, classés comme deux sous-espèces, l'une africaine (*L. siceraria* ssp. *siceraria*) et l'autre asiatique (*L. siceraria* ssp. *asiatica*), chacune ayant dû être domestiquée localement (Heiser 1973 cité dans Clarke et al. 2006). La gourde est retrouvée très tôt en Amérique, il y a 9 000 ans (Piperno et al. 2000). Il a été longtemps admis qu'elle avait été introduite naturellement (par flottaison) depuis l'Afrique (Heiser 1973 cité dans Clarke et al. 2006), résultat confirmé récemment par les données génétiques de gourdes américaines

**Tableau 1-1 : Liste des principales plantes du complexe horticole Lapitas**

Famille	Nom scientifique	Nom commun	Ancêtre sauvage	Foyer de domestication	Parties consommées	Parties utilisées pour la propagation	Système de reproduction	Ploïdie	Syndrome de domestication
<b>Racines et tubercules</b>									
Dioscoreaceae	<i>Dioscorea alata</i> L.	la grande igname	?	Nouvelle-Guinée?	tubercule	tubercule entier ou morceau de tubercule mais surtout entretenu et cueilli en forêt	dioïque	4x, 6x et 8x	?
Dioscoreaceae	<i>Dioscorea nummularia</i> Lam.	l'igname sauvage	même espèce	Domestication multilocale en Mélanésie?	tubercule	tubercule entier ou morceau de tubercule mais surtout entretenu et cueilli en forêt	dioïque	?	?
Dioscoreaceae	<i>Dioscorea bulbifera</i> L.	l'igname bulbifère	même espèce	Domestication multilocale en Mélanésie?	bulbille et parfois tubercule	bulbilles aériennes	dioïque	?	contre sélection de la toxicité, sélection des cultivars qui produisent des tubercules souterrains
Araceae	<i>Alocasia macrorrhiza</i> (L.) Schott	l'alocase ou oreille d'éléphant	même espèce	Asie du Sud-est? Nouvelle-Guinée?	corne	tête du corne ou rejet latéral	monoïque, allogame, fertilité réduite chez le cultivé	2x	?
Araceae	<i>Colocasia esculenta</i> (L.) Schott	le taro	même espèce	Asie du Sud-est et Nouvelle-Guinée	corne et feuille	tête du corne ou rejet latéral	monoïque, allogame (protogynie), fertilité réduite chez le cultivé notamment chez les triploïdes	2x et 3x	réduction de la quantité d'acide oxalique; augmentation de la taille du corne; raccourcissement des stolons
<b>Herbes</b>									
Musaceae	<i>Musa</i> sp.	le bananier	<i>Musa acuminata</i> ssp. <i>banksii</i> , puis hybridation en Asie du Sud-est avec d'autres <i>Musa acuminata</i> ssp. et <i>Musa balbisiana</i>	Nouvelle-Guinée	fruits	rejet	allogame mais sélection des cultivateurs pour des cultivars parthénocarpiques et stériles; stérilité des formes polyploïdes	2x, 3x et 4x	augmentation de la taille du fruit, contre sélection des graines
Poaceae	<i>Saccharum officinarum</i> L.	la canne à sucre	<i>Saccharum robustum</i> puis hybridation n Asie du Sud-est avec <i>S. spontaneum</i>	Nouvelle-Guinée	tige	bouture de tige immature	allogame (forte dépression de consanguinité)	polyploïde jusqu'à 10x	augmentation de la teneur en sucre
Agavaceae	<i>Cordyline fruticosa</i> L.	la cordyline	même espèce?	Asie du Sud-est ou Mélanésie?	feuille et aussi tubercule en Polynésie	bouture de tige	allogame; les formes de Polynésie orientale sont stériles	?	?

actuelles (Clarke et al. 2006). Pourtant, une étude moléculaire récente faite sur une dizaine d'échantillons archéologiques américains a démontré l'origine asiatique des restes archéologiques et conclu ainsi que cette plante a probablement été introduite, tout comme le chien, par les premiers migrants asiatiques qui ont peuplé l'Amérique par le détroit de Béring vers 15 000 ans avant JC (Erikson et al. 2005). Il est probable que des introductions européennes aient plus tard réintroduit la Gourde en Amérique depuis l'Afrique, où qu'elle soit aussi arrivée par flottaison, expliquant alors la discordance entre les données génétiques issues d'échantillons actuels et archéologiques (Clarke 2009). Des données linguistiques et archéologiques attestent de la présence de la Gourde dans tout le triangle polynésien dès 1200 après JC, alors qu'elle est absente des îles Fidji comme de Polynésie occidentale, mais bien présente en Asie du Sud-est et en Mélanésie (Clarke et al. 2006). Un tel patron de distribution suggère une origine duale : d'une part une origine asiatique à l'ouest du Pacifique, d'autre part une introduction en provenance de l'Amérique par flottaison, ou par les mouvements Polynésiens. Une récente étude génétique basée sur du matériel actuel retrace effectivement une double origine (origine africaine et asiatique) des variétés polynésiennes (variétés collectées uniquement en Nouvelle-Zélande) sur la base de marqueurs nucléaires, alors que les marqueurs chloroplastiques signent une origine principalement asiatique (Clarke et al. 2006). En l'absence d'échantillons archéologiques polynésiens et au vu de la discordance relevée précédemment, il semble difficile de conclure sur l'origine des allèles africains en Polynésie. Tout comme en Amérique, ils ont pu être introduits plus tardivement par les Européens (Clarke 2009).

La tomate polynésienne (*Solanum repandum*) a longtemps été considérée comme une introduction préhistorique américaine. Elle était largement distribuée des Marquises aux îles Fidji et toujours associée aux activités humaines. Ce patron de distribution suggère une introduction ancienne, probablement à partir d'une espèce sud-américaine similaire, *Solanum sessiliflorum*, les deux espèces étant probablement conspécifiques (Wistler 1991). La similarité linguistique entre un terme d'Amérique du Sud « *cocona* », et « *koko'u* » des îles Marquises serait aussi une indication supplémentaire d'un potentiel transfert préhistorique. Malheureusement, cette plante est aujourd'hui pratiquement absente des cultures polynésiennes (Wistler 1991).

Le cocotier aurait aussi pu traverser le Pacifique dans des temps préhistoriques avec les Polynésiens, comme en atteste l'existence de populations de cocotiers implantées avant l'arrivée des Européens sur les côtes Pacifique du Costa Rica, du Panama et de la Colombie (Green 2005). Toutefois, les dernières données génétiques montrent que ces cocotiers dérivent des Philippines directement et présentent peu de similitude génétique avec les autres variétés du Pacifique, notamment celles de Polynésie (Baudoin 2009, Gunn 2011). Ces résultats laissent plutôt ouverte la possibilité d'un transfert encore plus ancien par les austronésiens (il y a 2 250 ans) depuis les Philippines (Baudoin 2009).

La plante emblématique suggérant de tels contacts est sans conteste la patate douce (*Ipomoea batatas*). L'ensemble des données archéologiques, linguistiques et ethnobotaniques (cf. Chapitre 2) atteste qu'elle aurait été introduite en Polynésie dès 1100 ans après J. C. depuis les côtes d'Amérique du Sud (Nord du Pérou / Equateur), puis rapidement dispersée dans tout le triangle polynésien (Yen 1974, Green 2005). Les Polynésiens auraient pu poursuivre l'introduction de la patate douce plus à l'ouest, en Mélanésie, et peut être même jusqu'en Nouvelle-Guinée. Néanmoins, les données archéologiques et historiques sont ambiguës (Allen 2005, Ballard 2005 et Green 2005).

### *Les routes empruntées*

Depuis, l'analyse conjointe des courants, des données météorologiques et des systèmes de navigation polynésiens, couplée à des approches de simulation, ont largement démontré la possibilité de faire route depuis la Polynésie jusqu'en Amérique, puis d'en revenir (Irwin 2011). Le voyage aller nécessite de naviguer contre les vents et courants dominants venant d'est. Cependant, les Polynésiens auraient pu exploiter les inversions temporaires des vents dominants d'est en ouest, inversions annuelles régulières, mais aussi associées au phénomène météorologique El Niño. Trois connections sont



Tableau 1-1 (Suite)

Famille	Nom scientifique	Nom	Ancêtre sauvage	Foyer de	Parties	Parties utilisées pour la	Système de reproduction	Ploïdie	Syndrome de domestication
<b>Arbres, arbustes et arbrisseaux</b>									
Moraceae	<i>Artocarpus altilis</i> (Parkinson) Fosberg	l'arbre à pain	<i>A. camensis</i> puis hybridation avec <i>A. mariensis</i> en Micronésie	Nouvelle-Guinée	fruits	transplantation de drageons et graines	monoïque, allogame, (protandrie), possibilité de parthénocarpie, les triploïdes sont stériles	2x et 3x	augmentation de la taille du fruit, parthénocarpie, contre sélection des graines (stérilité)
Arecaceae	<i>Cocos nucifera</i> L.	le cocotier	même espèce	Deux foyers de domestication: l'Asie du Sud-est et le Sud du continent indien	noix	simplement cueillie ou plantation de la noix	monoïque, allogame mais possibilité d'autofécondation	2x	augmentation de la taille de la noix
Pandanaceae	<i>Pandanus</i> sp.	le pandanus	plusieurs espèces ( <i>P. antaresensis</i> , <i>P. brosimos</i> , <i>P. julianettii</i> , <i>P. tectorius</i> )	Domestication multilocale en Mélanésie	fruits et parfois feuilles	bouturage des rejets à la base des feuilles	?	?	augmentation de la taille du fruit?
Burseraceae	<i>Canarium indicum</i> et <i>Canarium harveyi</i>	la nangaille (noix de Kanari)	<i>Canarium indicum</i> et <i>Canarium Harveyi</i>	Domestication multilocale en Mélanésie	noix	aménagement des peuplements sauvages ou transplantation	?	?	augmentation de la taille de la noix
Piperaceae	<i>Piper methysticum</i> Forts. F.	le kava	<i>P. wichmannii</i> (probablement conspécifique)	Nord du Vanouatou	racine fraîche	bouture de tige	dioïque mais stérile	8x	sélection de chimiotypes particuliers
Metroxylaceae	<i>Metroxylon warburgii</i> (Heim.) Beccari	le sagou	même espèce	Nouvelle-Guinée	fécule du tronc	aménagement des peuplements sauvages ou transplantation	?	?	réduction des épines
Malvaceae	<i>Abelmoshus manihot</i> (L.) Medik	le chou des îles	même espèce	Asie du Sud-est? Nouvelle-Guinée?	feuille et fleur	bouture de tige	allogame mais possibilité d'autogamie, parfois stérile	polyploïde	?
<b>Introductions américaines</b>									
Convolvulaceae	<i>Ipomoea batatas</i> (L.) Lam.	la patate douce	<i>Ipomoea trifida</i> ?	Amerique tropicale	racine et parfois feuille	bouture de tige	monoïque, auto-incompatibilité sporophytique	6x et 4x?	formation de la racine tubéreuse, augmentation de la taille de la tige, prot rampant plutôt que grimpant
Cucurbitaceae	<i>Lagenaria siceraria</i> (Molina) Standl.	la gourde	<i>L. siceraria</i> ssp. <i>siceraria</i> et <i>L. siceraria</i> ssp. <i>asiatica</i> (trouvées uniquement à l'état cultivée)	Asie et Afrique indépendamment	fruit et feuille mais le fruit est surtout utilisé comme contenant	graine	monoïque, allogame mais possibilité d'auto-fécondation	2x	?

"?" signifie qu'il n'y a pas de données disponibles

aujourd'hui envisagées, en cohérence avec les autres preuves énoncées précédemment, connections qui ont pu avoir lieu une seule comme plusieurs fois indépendamment (Irwin 2011) (Figure 1-8).

\* La première route, la plus probable, est celle qui aurait introduit la patate douce en Polynésie : **la route Kumara**

Les polynésiens auraient ainsi fait route depuis l'Île de Pâques (des voyages en provenance de Tonga, des Samoa ou des Touamotou sont aussi possibles mais à certaines périodes de l'année seulement) jusqu'aux côtes du Pérou et de l'Equateur (100% de réussite à n'importe quelle période de l'année) vers 900 ans après J.C., où ils auraient récupéré la patate douce. La route de retour la plus probable les aurait conduits en Polynésie orientale entre 1000 et 1100 ans après J.C., quelque part entre les Marquises et les îles Cook.

\* Une autre route au nord, depuis Hawaii jusqu'aux côtes sud du Mexique et d'Amérique Centrale est aussi plausible si la navigation est effectuée entre Août et Septembre (entre 67 et 76% de réussite). Un voyage côtier jusqu'au sud de la Californie est aussi envisageable, le voyage de retour ayant pu profiter des alizés. La patate douce aurait ainsi pu être introduite à Hawaii au cours de tels voyages.

\* Il est enfin possible que des bateaux partant de l'Île de Pâques fassent route au sud et accostent sur les côtes chiliennes, entre Décembre et Janvier. Par contre, la route de retour jusqu'en Polynésie semble presque impossible, à moins d'avoir d'abord longé les côtes d'Amérique du Sud vers le nord, et ce jusqu'au nord du Pérou. Le poulet aurait ainsi pu être introduit au Chili au cours de tels voyages.

Un nombre grandissant de preuves encouragent donc à envisager sérieusement ces contacts préhistoriques entre Polynésiens et Américains (Jones et al. 2011), la patate douce restant sans doute l'exemple le plus suggestif. Peu d'études génétiques se sont pourtant emparées de cette énigme océanienne (Zhang et al. 2004, Clarke 2009), pourtant bien documentée par des données ethnobotaniques, linguistiques et archéologiques (Yen 1974, Green 2005). Le dernier chapitre de cette thèse (*chapitre 4*) s'attèle notamment à tester cette hypothèse d'introduction(s) préhistorique(s) de la patate douce depuis l'Amérique.

### ***II.2.7. Des plantes sur des pirogues***

La colonisation de l'Océanie lointaine s'est donc faite principalement de proche en proche, des plantes dans des pirogues, d'ouest en est pour la majorité des plantes, à l'exception notable de la patate douce introduite depuis l'est et diffusée à travers tout le triangle polynésien. Contrairement à des graines de céréales qui se conservent facilement et peuvent être emmenées en grande quantité, pour les plantes à multiplication végétative ce sont des boutures (rejets et têtes de cormes pour le taro, tubercules entiers ou morceaux de tubercules pour l'igname, morceau de tige ou racines pour la patate douce...), souvent encombrantes et périssables, qui ont été principalement transportées d'île en île (Lebot 2002). De forts goulots d'étranglement successifs au fil des migrations ont eu lieu, limitant la diversité à chaque nouvelle étape de colonisation. De plus, au cours de ces colonisations successives, les plantes ont aussi rencontré des conditions environnementales différentes, limitant parfois les possibilités d'adaptation.

#### *Un gradient de diversité d'Ouest en Est*

On observe un gradient de diversité des plantes cultivées d'ouest en est, tant au niveau spécifique qu'intra-spécifique (variétal) (Lebot 1992).

De plus d'une centaine d'espèces de plantes exploitées en Océanie proche et ce encore jusqu'au Vanouatou, environ 75 seulement sont introduites par les polynésiens à Tonga, autour de 45 dans les îles Cook, Sociétés et Marquises, une trentaine à Hawaii, entre six et huit sur l'Île de Pâques, et seulement six en Nouvelle Zélande (le taro, l'igname, le mûrier papier, la patate douce, la Gourde,

et la cordyline). Il est probable qu'un nombre plus important de plantes aient été introduites initialement en Nouvelle-Zélande, mais n'aient pas résisté au climat tempéré de ces îles (Wistler 1991, Kirch 2000).

Parallèlement, plusieurs études ont pu mettre en évidence un gradient de diversité génétique Ouest-Est au sein des cultivars d'une même espèce pour l'arbre à pain (Zerega et al. 2004), le taro (Lebot et Aradhya 1991, Lebot et al. 2004), la cordyline (Hinkle 2007). Ce gradient ne traduit pas nécessairement une baisse de la diversité des cultivars nommés, c'est-à-dire des variants morphologiques reconnus par les agriculteurs, mais bien le rétrécissement de la base génétique sous-jacente, c'est-à-dire une diminution de la diversité génotypique et allélique au fil des îles.

### *Des introductions multiples*

Des données de phylogéographie (basées sur l'analyse de l'ADN mitochondrial) obtenues à partir d'échantillons modernes et archéologiques de rat, *Rattus exulans* ont permis d'affiner notre image de l'odyssée polynésienne (Matisso-Smith et al. 1998). Elles démontrent notamment l'existence à Hawaïi, comme en Nouvelle-Zélande, de plusieurs voyages post-colonisation. En outre, les Polynésiens ont entrepris des voyages dits de retour vers l'ouest, peuplant ainsi en Mélanésie un certain nombre de petites îles (Kirch 2000). Ainsi, des réintroductions multiples de matériel végétal auraient pu limiter les effets de fondation des premières colonisations et réinjecter de la diversité cultivée au cours du temps.

### *Re-domestiquer localement certaines espèces*

Les migrants ont su profiter de la flore insulaire locale et certaines plantes ont probablement été domestiquées ou plutôt re-domestiquées localement dans différentes îles du Pacifique, limitant ainsi la perte de diversité génétique tout au long du processus de colonisation. Le genre *Pandanus* par exemple, exploité d'abord pour sa graine, mais aussi pour ses feuilles en vannerie (*P. tectorius* en Polynésie) est largement répandu de la Nouvelle Guinée à la Polynésie. Il est probable que différentes espèces aient été domestiquées plusieurs fois sur différentes îles (Yen 1995, Allaby 2007). De même, il est probable que l'igname *Dioscorea bulbifera*, présente à l'état spontané dans toute la Mélanésie ait elle aussi été re-domestiquée localement au cours des migrations. Ces possibilités de néo-domestication s'amenuisent au fur et à mesure que les migrants se déplacent vers l'est et s'installent sur les îles polynésiennes éloignées des continents et appauvries en espèces sauvages d'intérêt (Lebot 1992, 1999 et 2002).

### *Diffusion, propagation clonale et stérilité*

Un certain nombre de plantes océaniques présentent une fertilité réduite, certaines même sont stériles. L'apparition de la stérilité d'une plante à multiplication végétative au cours de sa diffusion peut avoir plusieurs origines distinctes.

Pour la banane, cette stérilité résulte d'hybridations interspécifiques en Asie du Sud-est, apparues très tôt dans le processus de diffusion (Perrier et al. 2011). Les variétés de bananiers diffusées dans le Pacifique appartiennent toutes au groupe « plantains du Pacifique » et correspondent à trois groupes variétaux morphologiques, très proches génétiquement, ayant chacun donné naissance à de nombreux mutants somatiques, sélectionnés par clonage (Lebot 2002). Ce même phénomène d'hybridation interspécifique serait à l'origine des formes stériles d'arbres à pain qu'on trouve

aujourd'hui en Micronésie. Là-bas, *Artocarpus altilis* se serait hybridé avec l'espèce locale *A. mariannensis* (Zerega et al. 2004).

Pour d'autres plantes, on note une apparition progressive de la stérilité au cours du processus de migration, mise en évidence notamment chez l'arbre à pain (Zerega et al. 2004), la cordyline (Hinkle 2007) ou encore le mûrier à papier (Seelenfreund 2010). L'arbre à pain est consommé principalement pour son fruit charnu riche en amidon, même si les graines sont aussi appréciées grillées (Walter et Lebot 2007). Les cultivars mélanésiens d'arbre à pain font des fruits à graines alors que les cultivars polynésiens sont aspermes. De même, les plantes de cordyline de Mélanésie jusqu'en Polynésie occidentale sont fertiles, alors que les formes de Polynésie orientale fleurissent mais semblent incapables de produire des graines. Parallèlement, on peut noter un changement d'usage de la cordyline d'ouest en est. En Mélanésie, outre ses nombreux usages culturels et religieux, elle est consommée essentiellement pour ses feuilles et peu pour ses racines, alors que les variétés polynésiennes possèdent un large rhizome apprécié (Hinkle 2007).

Ce gradient de stérilité est probablement à mettre en relation avec une sélection humaine accrue pour des formes « stériles », plus productives, au cours du processus de colonisation, sélection qui aurait été primordiale pour une conquête fructueuse des îles polynésiennes, lointaines et relativement pauvres en espèces utiles sauvages. En effet, le sexe à un coût : des génotypes qui investissent moins dans la reproduction sexuée et l'élaboration de fleurs et fruits, allouent probablement d'avantage de ressources à d'autres fonctions, conduisant souvent à une augmentation de la production d'autres organes végétatifs, comme les organes de stockage, par exemple (Mc Key et al. 2010). Ainsi, il existe un compromis entre la « fécondité clonale » et la fécondité sexuelle. Sélectionnant des génotypes très productifs via la propagation clonale, les agriculteurs pourraient par là même avoir contre-sélectionné la fécondité sexuelle de leurs plantes. Pour l'arbre à pain, cette stérilité est associée à un (ou plusieurs) événement de triploïdisation, qui aurait pu avoir lieu en Polynésie occidentale (Zerega et al. 2004). Ces clones triploïdes, particulièrement vigoureux, auraient alors été préférentiellement sélectionnés par les voyageurs polynésiens.

La baisse de la fertilité peut être aussi la conséquence d'une limitation des possibilités de croisements dans le cas d'espèces allogames obligatoires (Mc Key et al. 2010). Ainsi pour une plante dioïque (les ignames, ou le mûrier à papier) ou auto-incompatible (la patate douce), les effets de fondation successifs, lors du processus de migration, auraient introduit un nombre restreint de clones, conduisant à une stérilité « apparente » ou « écologique » : pour une plante dioïque, un des deux sexes pourrait avoir été majoritairement introduit dans certaines îles, limitant dès lors les possibilités de croisement. En Polynésie, il semble par exemple qu'on trouve essentiellement des arbres mâles de mûrier à papier (Wistler 1991, Seelenfreund 2010). De même, pour une plante auto-incompatible (comme la patate douce), l'introduction d'un nombre restreint de groupes d'auto-incompatibilité depuis l'Amérique aurait pu limiter les possibilités de recombinaison sur les îles polynésiennes.

En conséquence, il semble qu'en Polynésie, les agriculteurs gèrent des plantes qui ont une sexualité réduite, une base génétique réduite et dépendent plus fortement de la propagation clonale qu'en Mélanésie (Lebot 1992).

#### ***II.2.8. Pratiques paysannes et diversification des plantes à multiplication végétative dans les systèmes vivriers océaniques : rôles de la mutation somatique et de la reproduction sexuée***

Pourtant, les systèmes vivriers océaniques, polynésiens comme mélanésiens, sont caractérisés par une forte diversité variétale. Deux sources principales de diversification locale sont à la disposition des cultivateurs pour leurs plantes à multiplication végétative, la sélection de variants phénotypiques liés à l'apparition de mutations somatiques (modification de la séquence de la molécule d'ADN) ou de modifications épigénétiques transmissibles (modifications phénotypiques qui ne dépend pas d'un changement de la séquence de la molécule d'ADN mais par exemple d'une modification de l'état de condensation de la chromatine) et l'intégration de recombinants issus de graines.

Pour les plantes stériles (stérilité vraie comme stérilité apparente), la mutation somatique et les modifications épigénétiques restent les seules sources de variabilité phénotypique. Ces modifications sont susceptibles d'intervenir dans les zones méristématiques (méristèmes caulinaires ou racinaires), formant ainsi des chimères périclinales, où cohabitent, au sein du même organisme, des assises

cellulaires aux phénotypes distincts. Le panachage des plantes ornementales issu d'une mutation délétère sur la voie de synthèse de la chlorophylle est un exemple de mutations somatiques. Dans certains cas, ces cellules « mutantes » peuvent avoir un avantage sélectif sur celles des autres lignées, les remplacer progressivement, et ainsi se généraliser à l'organe mis en place. Par bouturage, les agriculteurs peuvent donc fixer ces phénotypes nouveaux dans de nouveaux clones. Même si les mécanismes précis de fixation chez les plantes à propagation clonale sont encore mal connus, il a été démontré chez de nombreuses plantes que ces mutations étaient une source de variabilité importante dans les systèmes océaniques, notamment en Polynésie. Ainsi, les nombreux morphotypes de bananiers, de taro ou de kava d'Hawaï présentent tous des zymotypes identiques (Lebot 1992), plus de 70 cultivars d'arbre à pain morphologiquement distincts décrits en Polynésie représentent des mutants somatiques d'un même clone initial (Ragone 1991, Zerega et al. 2007). De même, les taros polynésiens, pourtant très variables phénotypiquement, ont une base génétique très étroite et dérivent essentiellement par évolution clonale (Lebot et Aradhya 1991, Meilleur et al. 1998).

En Polynésie, les agriculteurs gèrent ainsi une grande diversité morphologique, mais une base génétique étroite pour la plupart de leurs plantes.

Plusieurs études ont aussi mis en évidence l'utilisation de la reproduction sexuée par les cultivateurs, surtout en Mélanésie. Ainsi, Caillon (2006) a montré que les agriculteurs des îles Banks au Nord du Vanouatou intégraient à leurs stocks de variétés des plantes issues de graines, trouvées après défrichement des anciens bassins à taros laissés en jachère pendant plusieurs dizaines d'années, rappelant ainsi les pratiques amérindiennes à l'égard du manioc (chapitre 1). Sardos et al. (2008) a aussi montré l'importance de la reproduction sexuée dans la diversification du manioc au Vanouatou. Il semble que la reproduction sexuée soit également impliquée dans la diversification de la grande igname au Vanouatou même si la fréquence de la production (et donc de l'incorporation) des graines semble moins élevée que pour le manioc ou encore le taro. Cette reproduction sexuée peut active peut s'expliquer par la dioecie (nécessité d'avoir les deux sexes à proximité et coordonnés), les différents niveaux de ploïdie et les frêles plantules issues de graines rapidement étouffées par les adventices (Sardos 2008). Pour la patate douce, Yen (1974) rapporte l'utilisation de plantules par les agriculteurs des Philippines et de Nouvelle-Guinée et propose même que ce soit le mode de diversification principale de la patate douce en Océanie. Dans les hautes terres de Nouvelle-Guinée, Schneider (1995) décrit l'intégration courante de plantules de patates douces dans les stocks de variétés. Le maintien de la reproduction sexuée dans ces systèmes hautement clonaux confèrent plusieurs avantages. Comme explicité dans la partie I), la reproduction sexuée permet d'éliminer les mutations délétères qui s'accumulent dans les lignées clonales (chapitre 1) ainsi que de combiner les différentes mutations favorables au sein d'un même individu. Elle confère donc une évolution plus rapide que dans un système strictement clonal. En outre, la propagation strictement clonale conduit souvent à une accumulation d'agents pathogènes. De nombreux agents pathogènes systémiques ne passent pas la barrière placentaire, les graines ne sont donc pas infectées et le passage par la reproduction sexuée permet alors d'assainir les plants au cours du temps (Mc Key 2010).

Les deux sources de diversification semblent donc se compléter et permettent aux agriculteurs de sélectionner de nombreux variants phénotypiques pour la plupart de leurs plantes océaniques. Cette diversité intra-spécifique a souvent été expliquée en termes écologico-fonctionnels, la sélection de multiples cultivars étant perçue comme une réponse adaptative des systèmes agricoles de ces îles vulnérables, exposées à des fortes variations environnementales et régulièrement soumises à d'importantes catastrophes naturelles (éruption volcanique, cyclone, tsunami...). Toutefois, l'étroitesse de la base génétique gérée par les agriculteurs laisse penser dans certains cas que les cultivars sont équivalents d'un point de vue fonctionnel, en particulier en Polynésie (Meilleur 1998). Dès lors, la diversité que l'on trouve sur ces îles océaniques répond tout autant à des exigences culturelles qu'écologiques. Elle est le reflet de la multitude d'usages culturels que les agriculteurs attribuent à leurs plantes (usages alimentaires, rituels, médicaux...), tout comme celles d'autres valeurs, certaines d'ordre patrimonial (support vivant de mémoire entre le passé et le présent) (Caillon et al. 2006), mais aussi cosmologique ou encore esthétique (Meilleur et al. 1998).

*« La coutume a d'abord cherché à répondre à ce problème fondamental : comment peut-on vivre dans une île ? Posée de cette façon, la question renvoie aux prémices : l'identité mélanésienne s'est forgée dans un monde où il semble que son premier problème ait été de survivre, non pas*

*matériellement, mais culturellement. Dans cet univers clos, elle a cherché à recréer des conditions maximales de diversité culturelles et sociales. En donnant un sens au paysage, elle a redonné un sens à la vie de ceux qui l'habitent* », extrait du livre « les gens des lieux » de J. Bonnemaïson (1997).

## **II.2.9. L'arrivée puis l'installation des Européens**

### *Introduction et re-distribution des plantes cultivées*

L'installation des Européens dans le Pacifique est constituée de plusieurs phases, correspondant à une intensification progressive des mouvements humains et par là même une intensification des échanges, notamment d'espèces végétales.

La première période correspond à des expéditions exploratoires et scientifiques. Après la traversée du Pacifique par Magellan (1520-1521) suivent les deux expéditions de Mendaña, l'une en 1558 où il découvre les îles Salomon et les Santa Cruz puis en 1595 les îles Marquises. Queiros, en 1606 accostera au Touamotou ainsi que sur l'île de Santo (Vanouatou). Les autres voyages d'exploration ne reprendront qu'à partir du milieu du 18<sup>ième</sup> siècle. Parmi eux, le voyage de *l'Etoile*, commandée par le capitaine Antoine de Bougainville qui explorera l'archipel des Touamotou, Tahiti ainsi que les îles Samoa. Il sera suivi par les trois expéditions successives du Capitaine James Cook (de 1768 à 1780), qui accosteront notamment sur l'Île de Pâques, Tahiti, Tonga, la Nouvelle-Zélande et Hawaïi. Cette tradition d'exploration continuera après Cook jusqu'au milieu du 19<sup>ième</sup> siècle avec entre autres les voyages de La Pérouse, D'Entrecasteaux, Malaspina et Dumont D'Urville. Ces voyages à vocation exploratoire étaient surtout destinés à ramener des plantes en Europe, mais il est avéré qu'ils ont très tôt contribué à l'introduction de plantes cultivées dans le Pacifique. Ainsi Queiros planta du maïs au Marquises et son livre de bord indique clairement qu'il planta des « patates » dans un petit jardin à Matantas, dans la grande baie de Santo (Baert 2001). Cook introduisit des *Citrus* sp. et l'ananas à Tahiti (Barrau 1962).

Dès 1565, les espagnols établirent une connexion commerciale régulière entre le Mexique et les Philippines qui dura jusqu'en 1815. Les galions assurant cette connexion transportaient principalement des marchandises et contribuèrent à l'introduction précoce de nombreuses plantes alimentaires américaines en Asie du Sud-est, plantes qui seront ensuite distribuées en Océanie par les migrants locaux ou plus tard par les Européens (l'ananas, l'anone *Annona reticulata* entre autres). La patate douce mexicaine appelé « Camote » (en langue Nahuatl) est alors introduite aux Philippines dès le 16<sup>ième</sup> siècle (Barrau 1957, Yen 1974). A la même époque, les Portugais eux aussi déplacent des plantes de leurs colonies américaines jusqu'à leurs comptoirs indonésiens, la patate douce entrant ainsi en Indonésie depuis les caraïbes par la route portugaise dite « Batata » (Barrau 1957, Yen 1974).

Durant la seconde moitié du 19<sup>ième</sup> siècle, des sociétés missionnaires et autres colons viennent s'installer sur les îles du Pacifique et vont jouer un rôle prépondérant dans l'introduction de nouvelles espèces. On peut citer l'exemple du naturaliste russe Mikloukho Maklai, qui explore la côte nord de la Nouvelle-Guinée entre 1871 et 1879 et y introduit des plantes comme la papaye ou encore le maïs et la pastèque, connus des autochtones sous leurs noms russes (Barrau 1962). Par la suite, les plus grandes facilités de communication qui résultèrent de cette présence européenne en Océanie facilitèrent les échanges de plantes entre les archipels restés jusqu'à là relativement isolés les uns des autres. Ces mouvements vont permettre dans un premier temps de redistribuer la diversité végétale élaborée localement. Lorsque les insulaires réussissent à maintenir un équilibre entre ces apports extérieurs et leurs économie vivrières, ces introductions vont même participer à ré-enrichir la diversité cultivée. Pourtant dans d'autres cas, l'influence européenne va bel et bien déstabiliser cette économie vivrière entraînant même parfois son déclin (Barrau 1962).

Le Pacifique jusqu'alors vaste espace ouvert à ses seuls habitants, reliés entre eux par de rares contacts et de frêles embarcations devient un espace mondialisé. Il est d'abord le siège d'une colonisation forcenée, puis est rapidement propulsé dans la modernité et l'économie de marché capitaliste (Antheaume et Lawrence 1985). L'implantation de cultures commerciales de cocotiers (pour l'exploitation du coprah), de canne à sucre, de cacao ou encore de café commence dès le milieu du 19<sup>ème</sup>. Ces grandes plantations européennes occupent souvent de larges espaces de terres fertiles empiétant sur les cultures traditionnelles (Yen 1998). Plus tard, des programmes de développement inciteront les agriculteurs à amorcer une transition vers une agriculture commerciale (Antheaume et Lawrence 1985). L'essor de cette agriculture commerciale, souvent de plantes pérennes, va immobiliser durant de longues années des territoires auparavant disponibles pour les rotations de cultures (Lebot 2002). Dès lors, cette agriculture commerciale entraîne de fortes pressions sur les terres cultivables, une intensification de l'utilisation des terres restantes, un accroissement de la déforestation et un raccourcissement des temps de jachères. Ces transformations agricoles induisent en général une baisse de la performance des systèmes vivriers traditionnels (Yen 1998 ; Siméoni et Lebot 2012). En outre, le développement économique des îles (urbanisation, industrialisation, développement du tourisme...) permet aux insulaires d'accéder à de nouvelles sources de revenus. Ils délaissent alors leurs pratiques vivrières et leurs plantes traditionnelles au profit d'autres activités et de produits alimentaires importés (Barrau 1962, Yen 1998). Sinon, les activités vivrières se concentrent sur l'utilisation de certaines plantes alimentaires exotiques comme le macabo, le manioc ou encore la patate douce qui permettent d'assurer une production alimentaire de base pour un travail agricole réduit (Walter et Lebot 2007). Dans ces îles, l'activité vivrière est plus qu'une simple économie mais embrasse littéralement toutes les dimensions de la vie sociale (Barrau 1962, Bonnemaïson 1996). L'érosion culturelle de ces sociétés compromet nécessairement le maintien de l'agriculture traditionnelle et vice et versa. Dès lors, la diversité des plantes cultivées qui lui est intimement liée, est elle aussi menacée.

Les voyageurs européens ont très tôt introduit de nombreuses maladies sur les îles du Pacifique décimant ainsi les populations insulaires. A Guam par exemple, la population est passée de 40 000 à 1 500 entre 1668 et 1720. Cette chute brutale de la population aurait d'ailleurs favorisé l'installation des européens et la réquisition de nombreuses terres pour l'installation des grandes plantations et pâturages (Nunn 2003). L'intensification des flux migratoires et des échanges inter-insulaires depuis la seconde guerre mondiale (commerces, tourisimes) a aussi conduit à l'introduction de nombreux agents pathogènes et prédateurs des cultures traditionnelles. Ainsi, l'introduction en Polynésie du *Phytophthora colocasiae* depuis la Nouvelle-Guinée et les îles Salomon a ravagé il y a vingt ans les cultures de taro des îles Samoa (Lebot 2002). La base génétique assez étroite des cultures océaniques dans certaines régions limite dès lors les possibilités de faire face à ces nouvelles maladies et pose de véritables problèmes phytosanitaires.

Les îles du Pacifique connaissent aujourd'hui une forte pression démographique (de l'ordre de 2% en Mélanésie). La population actuelle dépasse les 10 millions d'habitants et devrait atteindre les 15 millions en 2035 ([www.spc.int/](http://www.spc.int/)). De plus, ces îles subissent de plein fouet les conséquences des changements climatiques en cours (Nunn 2009). Dans ces îles, on s'attend à une augmentation de la variabilité du climat, de la fréquence des cyclones et des périodes de sécheresse dues aux cycles el Niño - la Niña et à un accroissement de la température moyenne et du niveau moyen des mers. Les modèles prédisent que ces variations environnementales devraient diminuer la productivité des systèmes de culture (Nunn 2009). Ce n'est pourtant pas la première fois que les populations Océaniques doivent faire face à de rapides changements environnementaux au cours de leur histoire (Nunn 2000). En effet, il y a environ 700 ans, de brusques changements ont eu lieu (refroidissement des températures et baisse du niveau de la mer), provoquant de profondes transformations au sein des sociétés océaniques. En outre, il est notable que certaines îles ont connu par le passé des densités de populations bien plus élevées que celles rencontrées aujourd'hui (Kirch et Rallu 2007). Au Vanouatou par exemple, la population s'élevait en 2009 à 234 000 habitants, alors qu'en 1800 elle comptait probablement entre 500 000 et 1 million d'habitants (Siméoni et Lebot 2012). Dans ces contextes de forte densité de population, les systèmes vivriers se sont avérés performants et leur forte résilience a

jusqu'à présent permis aux Océaniens de s'adapter aux variations de leur milieu (Nunn 2003, Siméoni et Lebot 2012). On peut anticiper que ces systèmes vivriers traditionnels peuvent encore aujourd'hui jouer ce rôle et assurer l'autonomie alimentaire des îles du Pacifique, à condition qu'ils soient revalorisés. Cela implique, entre autres mesures, une re-valorisation des ressources génétiques des espèces traditionnelles et la recherche de cultivars adaptés aux besoins locaux et aux contraintes agronomiques nouvelles (pressions pathogènes, productivité, résistance à la sécheresse...) (Lebot 2002).

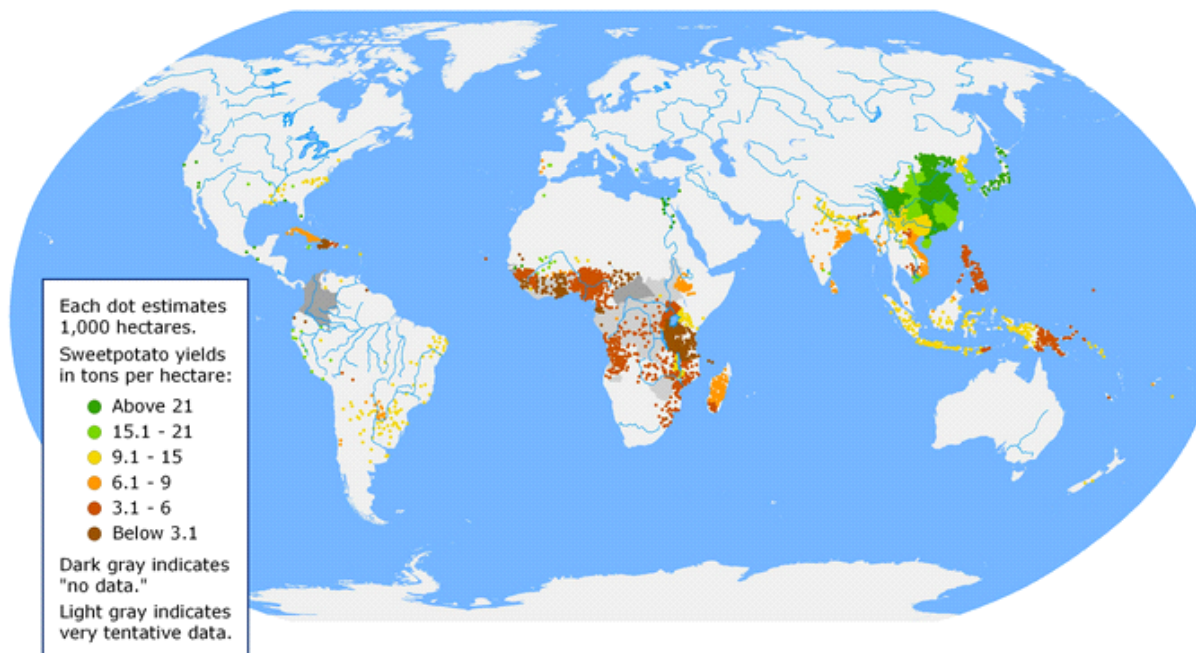
Dans cette thèse, nous nous intéressons à l'une de ces cultures Océaniques, la patate douce. Originale d'Amérique tropicale, elle a été introduite en Océanie plus tardivement que le taro ou l'igname, néanmoins - dans certaines régions- avant l'arrivée des Européens. Par le passé, elle a joué un rôle déterminant dans l'adaptation des sociétés océaniques (chapitre 2) et devrait continuer à y tenir une place primordiale. Bien que l'importance alimentaire de cette plante soit largement reconnue, il est étonnant de voir que son histoire évolutive, l'histoire de sa domestication (chapitre 3 et 4) et de sa diffusion en Océanie (Chapitre 5 et 6) ainsi que les mécanismes qui ont façonné cette diversité au cours du temps et dans l'espace ont été peu étudiés. Ce travail de thèse se propose d'éclairer quelques aspects de cette histoire.



# *CHAPITRE 2*

*Modèle biologique, problématiques et outils de  
l'enquête*

**Figure 2-1 : Aperçu de la production mondiale de patate douce.** (d'après l'atlas mondial de la patate douce, Centre international de la Pomme de terre, Lima)



© <https://research.cip.cgiar.org/confluence/display/WSA/Global+Sweetpotato+Cultivation>

La patate douce (*Ipomoea batatas* (L.) Lam., Convolvulaceae) est classée au septième rang dans les statistiques alimentaires mondiales (Lebot 2009). Elle est essentiellement consommée pour sa racine tubéreuse, qui possède une forte valeur nutritionnelle (source importante d'amidon et de vitamines). Les variétés orange sont une excellente source de bêta-carotène, un précurseur essentiel de la synthèse de la vitamine A. Ses feuilles sont riches en protéines. C'est une plante rustique qui possède une grande adaptabilité agronomique et tolère une large gamme de conditions édaphiques et climatiques. Cultivée préférentiellement en conditions tropicales, elle résiste particulièrement bien au froid et est ainsi cultivée dans de nombreuses régions de hautes-terres en Afrique de l'Est ainsi qu'en Nouvelle-Guinée. Elle peut produire plus d'énergie consommable par jour que d'autres plantes alimentaires comme le blé, le riz, ou encore le manioc (Lebot 2009). Aussi, c'est une plante relativement facile à cultiver qui sous les tropiques ne montre pas de caractère saisonnier strict, a un cycle relativement court (environ quatre mois) et nécessite peu d'intrants. Elle est généralement consommée grillée, bouillie ou encore frite, mais est aussi largement utilisée dans l'industrie comme source d'amidon pour la confection de produits dérivés (nouilles, sucreries, pectines, alcool). Elle est aussi destinée à l'alimentation animale.

La patate douce est produite dans plus d'une centaine de pays (Figure 2-1). Bien que cette plante soit originaire d'Amérique tropicale, l'Asie est aujourd'hui le plus gros producteur mondial (plus de 125 millions de tonnes de racines fraîches par an), la Chine produisant à elle seule plus de 100 millions de tonnes. En Amérique tropicale, elle est surtout cultivée dans les régions très pauvres, à Haïti, ou encore à Cuba. En Afrique, c'est une culture d'« hommes pauvres », produite essentiellement dans des exploitations familiales. Toutefois de nombreux projets de développement agricole cherchent à en promouvoir la culture afin de lutter contre la malnutrition et les carences en vitamines A. En Océanie, la patate douce est aujourd'hui la plante alimentaire principale des habitants de Nouvelle-Guinée et des îles Salomon, où elle représente plus de 66% de l'énergie alimentaire produite localement. En Papouasie Nouvelle Guinée, on évalue à 685 kg par an et par personne, la consommation de tubercules en zone rurale (Bourke 2009). Malgré son importance alimentaire reconnue, son histoire évolutive est encore mal connue, comparativement à d'autres plantes cultivées.

## I. ORIGINES ET DOMESTICATION DE LA PATATE DOUCE : FAITS ET HYPOTHESES

### I.1. Position taxonomique

La patate douce a tout d'abord été décrite par Carl Von Linné en 1753 comme *Convolvulus Batatas*. Cependant, dès 1791, Lamarck re-classifia cette espèce à l'intérieur du genre *Ipomoea* (sur la base de la forme des stigmates et la surface des grains de pollen). Le nom actuel est *Ipomoea batatas* (L.) Lam.

**Famille:** *Convolvulaceae*

**Tribu:** *Ipomoeae*

**Genre:** *Ipomoea*

**Sous-genre:** *Eriospermum*

**Section:** *Eriospermum*

**Serie:** *Batatas*

**Espèce:** *Ipomoea batatas* (L.) Lam.

#### I.1.1. Le genre *Ipomoea*

Le genre *Ipomoea* est le plus vaste de la famille des Convolvulacées (une des grandes familles, avec celle des Solanacées, de l'ordre des Solanales). Il compte entre 600 et 700 espèces réparties à travers le monde dans les régions tropicales et subtropicales et occupant des habitats très variés (Austin 1996).

**Tableau 2-1 : Les 14 espèces d'*Ipomoea* de la série *Batatas*, leur niveau de ploïdie et leur aire de répartition.**

<b>Espèces</b>	<b>Ploïdie x=15</b>	<b>Aire de répartition</b>
<i>I. cyanchifolia</i>	2x	Amérique du Sud
<i>I. lacunosa</i>	2x	Amérique du Nord
<i>I. leucantha</i>	2x	Amérique du Nord
<i>I. littoralis</i>	2x	Pacifique et Océan Indien
<i>I. ramosissima</i>	2x	Amérique centrale et du Sud
<i>I. tenuissima</i>	2x	Iles Caraïbes et Amérique du Nord
<i>I. triloba</i>	2x	Amérique centrale et du Sud
<i>I. umbraticola</i>	2x	Amérique centrale
<i>I. tiliacea</i>	4x	Amérique centrale et îles Caraïbes
<i>I. tabascana</i>	4x	Mexique
<i>I. cordotriloba</i> ( <i>I. trichocarpa</i> )	2x, 4x	Amérique centrale et du Nord
<i>I. grandifolia</i>	2x	Amérique du Sud
<i>I. trifida</i>	2x, 3x?, 4x?, 6x?	Amérique centrale et du Sud
<i>I. batatas</i>	4x?, 6x	Amérique tropicale

Plus de la moitié des espèces se concentrent sur le continent américain. La plupart des Ipomées sont des plantes herbacées, annuelles ou pérennes, volubiles ou rampantes, certaines pouvant aussi avoir un port arbustif. Ce genre regroupe la plupart des espèces utiles et d'importance économique de la famille des Convolvulacées.

### ***1.1.2. Le complexe ou série Batatas***

Le genre *Ipomoea* est subdivisé en trois sous-genres, subdivisés en sections puis en séries. La série *Batatas*, d'abord délimitée par Van Oostrom (1953), puis Vercourt (1963), a été finalement révisée par Austin (1978). Cette section comporte aujourd'hui 14 taxons (incluant la patate douce), représentés par quelques plantes annuelles et une majorité de plantes pérennes (Tableau 2-1 et Figure 2-2). La plupart des espèces se rencontrent dans les milieux ouverts des basses terres d'Amérique tropicale, à l'exception notable d'*I. littoralis*, une espèce native et endémique d'Asie tropicale qu'on rencontre principalement sur les côtes et dans les îles du Pacifique et de l'Océan indien. Trois taxons dériveraient d'hybridations interspécifiques : *I. leucantha*, *I. grandifolia* et *I. tabascana* (Austin 1977). La distinction morphologique entre ces espèces, ainsi que l'étude de leurs relations phylogénétiques, est complexe, en raison d'une grande variabilité intra-spécifique de la plupart des traits, d'éventuels phénomènes d'homoplasie associés, ainsi que de l'existence d'hybridations interspécifiques (Austin 1988). Ceci cause d'ailleurs de nombreux problèmes d'identification du matériel végétal, qui ont des conséquences importantes dans l'interprétation des relations phylogénétiques entre espèces (D.F. Austin, comm. pers. 2012). La forme et l'aspect des sépales restent toutefois des critères pertinents pour l'identification des différentes espèces (Austin 1988) mais leur étude requiert souvent l'œil avisé d'un taxonomiste (D.F. Austin, comm. pers. 2012). Deux sous-groupes (A et B) sont caractérisés sur la base de données morphologiques et cytogénétiques, ainsi que l'analyse des données de croisements interspécifiques (Jones 1965 ; Jones et Deonier 1965 ; Martin et Jones 1973 ; Nishiyama et Sakomoto 1972 ; Austin 1988 ; Oracion et al. 1990 ; Srisuwan et al. 2006). Plus récemment, des études moléculaires ayant recours à différents types de marqueurs (RFLP, RAPD, ISSR et enfin un gène nucléaire) ont confirmé l'existence de ces deux sous-groupes, même si *I. littoralis* et *I. tiliacea* ont une position phylogénétique parfois changeante selon l'étude (Jarret et al. 1992 ; Jarret et Austin 1994 ; Buteller et al. 1999 ; Huang et Sun 2000 ; Huang et al. 2002 ; Rajapakse et al. 2004 ; Gao et al. 2011). En résumé, le groupe B, monophylétique, comporte les espèces *I. batatas*, *I. trifida* et *I. tabascana*, un hybride potentiel entre les deux taxons précédents (McDonald et Austin 1990). Toutes sont auto-incompatibles et s'hybrident naturellement. *Ipomoea littoralis* est parfois associé à ce groupe génomique B (Rajapakse et al. 2004). Le groupe A est représenté par l'ensemble des autres espèces de la série (hormis *I. umbraticola* et *I. ramossissima*, qui sont plus distantes), principalement auto-compatibles (à l'exception de *I. tiliacea*), certaines inter-fertiles. L'hybridation des espèces de ces deux sous-groupes est rare mais possible (Figure 2-3) (Diaz et al. 1996).

Différents niveaux de ploïdie sont reportés dans cette section. La plupart des espèces sont diploïdes ( $2x = 30$ ), certaines exclusivement polyploïdes comme *I. tiliacea* ( $4x = 60$ ) ou *I. batatas* ( $4x = 60$  et  $6x = 90$ ). *Ipomoea trifida* est défini comme un complexe de lignées conspécifiques, adaptées à différentes conditions écologiques (écotypes) et avec différents niveaux de ploïdie ( $2x$ ,  $3x$ ,  $4x$  et  $6x$ ) (Kobayashi 1984).

## **1.2. Aux origines de *I. batatas* : retour sur une controverse**

La question des origines de la patate douce pose en fait une double question, celle de l'origine de l'espèce *Ipomoea batatas* et celle de l'origine de la forme cultivée. Ces deux questions sont souvent confondues. La difficulté de répondre à cette double interrogation trouve son origine dans une confusion taxonomique qui a perduré longtemps et s'est immiscée dans la plupart des études génétiques et/ou cytogénétiques qui se sont penchées sur l'origine de *I. batatas* (D.F. Austin, 2012).

D'un point de vue taxonomique, l'espèce *I. batatas* (D. F. Austin, 2012) est l'unique espèce hexaploïde de la série *Batatas* et la seule capable de produire de véritables racines tubéreuses comestibles (Austin 1988). Des populations sauvages de *Ipomoea* principalement tetraploïdes mais

Figure 2-2 : Espèces d'*Ipomoea* mentionnées dans cette étude (Chapitre 4).



d'après Srisuwan 2006



aussi quelques formes hexaploïdes, ont été collectées du Mexique au Nord du Pérou (Jones 1967 ; Nishiyama 1972 ; Kobayashi 1984 ; Austin 1988 ; Bohac et al. 1993). Ces plantes sont proches de *I. batatas*, mais apparaissent régulièrement dans la littérature sous diverses appellations, parfois sous le nom de *I. littoralis* ou *I. gracillis*, et très souvent sous le nom de *I. trifida* (Shiotani 1990 ; Kobayashi 1984 ; Bohac et al. 1993). Elles sont décrites comme des formes sauvages car elles ne produisent pas de vraies racines tubéreuses mais seulement des racines épaissies. Des travaux menés sur la base de critères botaniques (taille et forme des sépales) ont permis de ré-identifier la plupart de ces accessions comme des formes intermédiaires, sauvages et/ou échappées de culture, de *I. batatas* (Bohac et al. 1993). Des doutes ont alors été émis quant à l'existence de véritables formes d' *I. trifida* polyploïdes (Austin 1988). Rares sont pourtant les travaux de génétique et de cytogénétique qui ont intégré ces formes intermédiaires et elles sont la plupart du temps considérées comme des *I. trifida* (Komaki et al. 1998 ; Rakjapakse et al. 2004 ; Srisuwan et al. 2006). Elles sont donc mal caractérisées d'un point de vue génétique. Elles représentent pourtant un maillon essentiel dans la reconstruction de l'histoire évolutive de *I. batatas* (Bohac et al. 1993). Malheureusement, la distribution, tout comme l'écologie de ces formes sauvages, est peu documentée. Il en existe cependant quelques représentants dans certaines collections *ex situ*, souvent cachées sous d'autres appellations (Figure 2-2) (Kobayashi 1984 ; Bohac et al. 1993 ; Komaki et al. 1998).

### I.3. L'origine botanique de *I. batatas* : deux scénarios en jeu

Deux scénarios s'affrontent pour expliquer l'origine du génome hexaploïde de *I. batatas* (Figure 2-4). Selon certains auteurs la patate douce a une origine hybride, impliquant au moins deux génomes distincts. L'analyse des caractères morphologiques amena Austin (1988) à conclure que *I. triloba* et *I. trifida* devaient être les parents sauvages. Ce scénario allo-autopolyploïde est conforme à des études anciennes de cytogénétique sur la méiose de *I. batatas*. Les associations secondaires entre bivalents indiquent que le génome de *I. batatas* est composé de deux génomes homologues et d'un autre plus distant (Ting et Kehr 1953, et Jones 1965). La ségrégation tétrasomique révélée sur la base de marqueurs microsatellites est cohérente avec cette hypothèse (Buteller et al. 1999). Toutefois, l'implication de *I. triloba* n'a trouvé à ce jour aucune validation moléculaire alors que toutes les études de phylogénie et de cytogénétique moléculaires citées précédemment ont établi la proximité génétique de *I. batatas* à *I. trifida*. Cependant, les techniques moléculaires employées jusqu'ici ne permettent pas d'étudier la composition réelle des six copies potentielles du génome hexaploïde (copies homologues comme homéologues) et de résoudre « proprement » la question de l'origine phylogénétique de la patate douce.

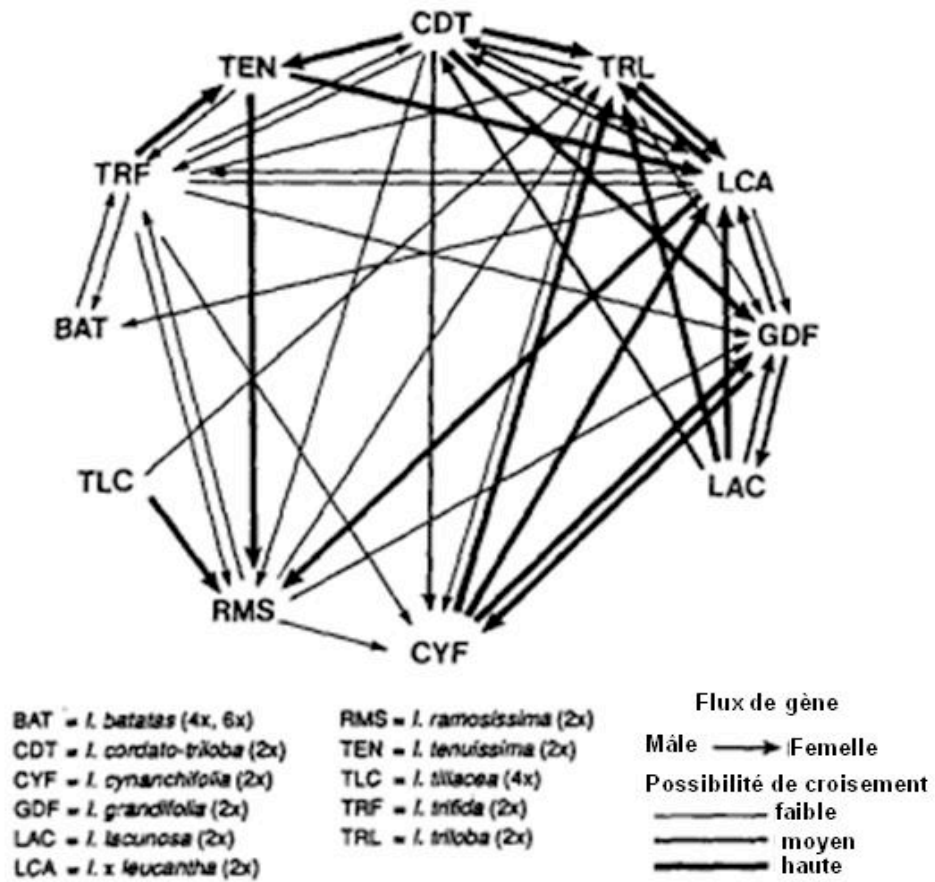
D'autres auteurs suggèrent que la patate douce cultivée dérive de populations de *I. trifida* sauvages diploïdes par autopolyploïdisations successives (Figure 2-4). Kobayashi (1984) introduit l'idée du « complexe *I. trifida* » qui comprend l'ensemble des formes diploïdes et polyploïdes de la série *Batatas*, se croisant avec la patate douce. Selon cette hypothèse, la forme cultivée dériverait de la domestication de formes hexaploïdes d'*I. trifida*. Les travaux récents vont dans le sens de cette dernière hypothèse. Cependant, considérant la confusion taxonomique explicitée plus haut, la notion de « complexe *I. trifida* » doit être revue dans la mesure où la plupart des accessions polyploïdes qui ont servi à élaborer ce concept sont très probablement des *I. batatas*.

La controverse est loin d'être close. Une récente étude basée sur l'analyse des variants de l'intron 2 du gène *waxy* (Gao et al. 2011) conclue à une origine allopolyloïde de *I. batatas*. Elle serait issue de l'hybridation entre *I. littoralis* et *I. tenuississima*. Malheureusement, les études génétiques menées à ce jour se sont souvent focalisées sur les relations phylogénétiques entre espèces et ont délaissé des approches phylogéographiques et populationnelles, pourtant nécessaires à l'étude d'un groupe taxonomique sujet à l'hybridation et à la polyploïdisation.

### I.4. Origine géographique de la domestication

Il est admis que la patate douce est originaire d'Amérique tropicale (Austin 1988). Considérant la distribution actuelle des espèces sauvages apparentées, Austin (1988) a proposé que l'ancêtre sauvage

Figure 2-3 : Possibilités d'hybridation entre les différentes espèces du complexe *Batatas*.



(d'après Diaz et al. 1994)



de *I. batatas* ait été formé quelque part entre la péninsule du Yucatan au Mexique et le bassin de l'Orénoque au Venezuela. Les plus vieux restes archéologiques de patate douce ont pourtant été retrouvés dans différents sites archéologiques de la côte péruvienne (Ugent et Peterson 1998), les plus anciens pouvant même dater de 10 000 ans (Engel 1970). L'étude de la structure de la diversité génétique des cultivars d'Amérique tropicale sur la base de marqueurs AFLP (Zhang et al. 2000) et RAPD (Gichuki et al. 2003) suggère l'existence de deux groupes génétiques différenciés, le groupe mésoaméricain (le plus divers) et le groupe péruvo-équatorien. Cette différenciation est interprétée comme le résultat d'un foyer de domestication unique en Amérique Centrale suivi de la diffusion rapide de la plante jusqu'en Amérique du Sud, puis d'une diversification secondaire dans cette région (Zhang et al. 2000). Ce patron de diversité pourrait tout aussi bien être issu d'événements de domestication indépendants, scénario sous-tendu par l'ancienneté des restes archéologiques sur la côte péruvienne. Cependant, cette hypothèse a rarement été envisagée dans la littérature. Pour cause, aucune donnée morphologique ne vient supporter l'existence de deux foyers de domestication, contrairement à d'autres espèces comme le haricot, qui ont une origine multiple et où des races morphologiques sont identifiables (Chacón et al. 2005).

## **I.5. Le caractère domestiqué : la racine tubéreuse**

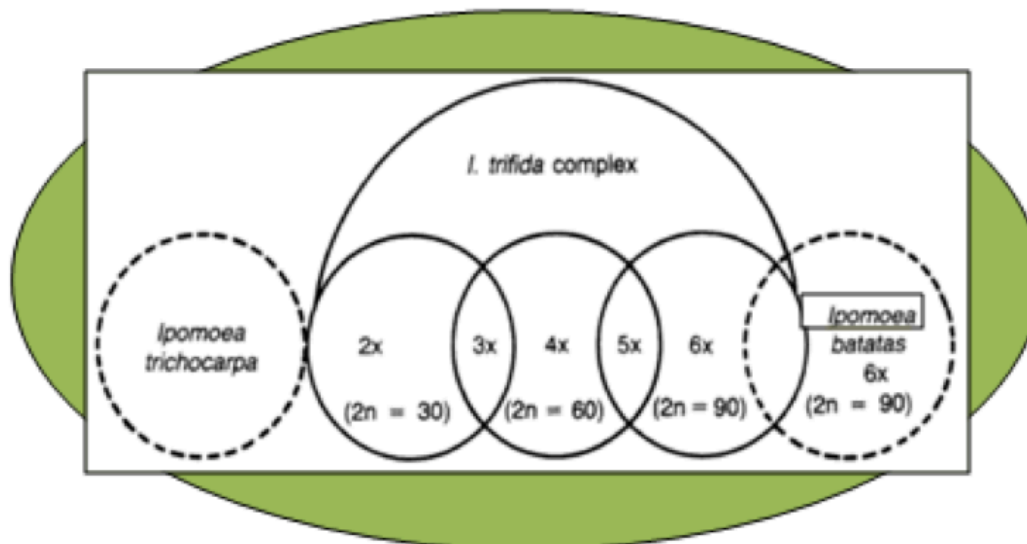
Le syndrome de domestication de la patate douce est mal délimité puisque son ancêtre sauvage n'est toujours pas clairement identifié. Il semble que la domestication ait principalement porté sur la sélection de plantes produisant des racines tubéreuses comestibles, qui se développent par croissance secondaire des racines adventives. Toutefois, on s'attend aussi à une augmentation de la taille des tiges qui servent de boutures, comme chez le manioc (Rival et McKey 2008) et probablement aussi à des modifications du port, très volubile chez les espèces sauvages apparentées et souvent rampant chez la patate douce cultivée. Les autres espèces de la série *Batatas* portent parfois des racines latérales légèrement épaissies, notamment *I. trifida*, *I. ramosissima* et *I. tiliacea* (Diaz et al. 1992), mais sans intérêt alimentaire attesté. Aussi, ce caractère principal serait lié à l'état hexaploïde puisqu'aucun des spécimens tétraploïdes (*Ipomoea* spp. identifiées comme *I. batatas*) collectés ne produit de véritables racines tubéreuses (Bohac et al. 1993). Ce caractère est néanmoins instable au sein même du compartiment cultivé. Il dépend des conditions du milieu : certains génotypes font des racines tubéreuses dans certaines conditions mais pas dans d'autres. Par exemple, les variétés des hautes terres de Nouvelle-Guinée ne forment pas de racines tubéreuses lorsqu'elles sont plantées dans les basses terres (Van Wijmeersch 2001). Aussi, à l'issue d'un événement de recombinaison entre deux variétés (même si elles font toutes deux de beaux tubercules), une certaine fraction des plantules ne font pas de racines (Jones 1967). Le passage par la propagation clonale permettrait alors de stabiliser ce caractère. Les bases génétiques de ce trait ne sont pas connues précisément.

## **I.6. Biologie de la reproduction et sources de variabilité**

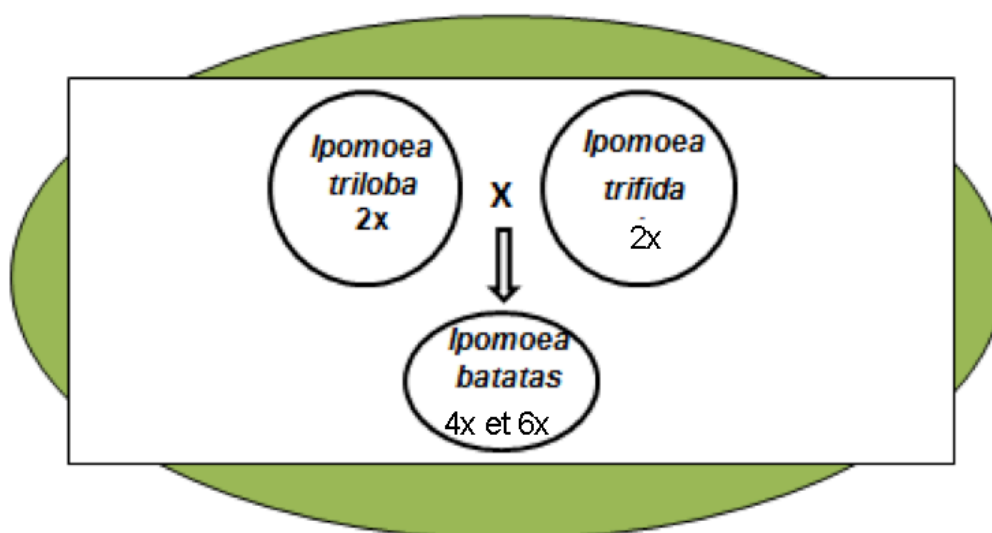
La patate douce est multipliée végétativement par les agriculteurs, principalement par l'usage de boutures de tige mais aussi parfois de racines de réserves entières, qui sont munies de bourgeons à leur surface (Yen 1974). La floraison est un phénomène relativement courant en conditions tropicales (Yen 1974 ; Lebot 2009) (Figure 2-5a), plus rare en milieu équatorial. En effet, la formation de la fleur est contrôlée par l'environnement et dépend de la photopériode (induite par le raccourcissement des jours), de la température (induite notamment par le froid) et aussi du génotype de la plante (Yen 1974). Les fleurs s'ouvrent le matin très tôt, restent ouvertes pour quelques heures puis se ferment et fanent avant midi dans la même journée. Elles sont donc réceptives pendant une courte période. La pollinisation est entomophile, en particulier par les abeilles. La fécondation aboutit à la formation de fruits qui sont des capsules contenant une à trois graines chacune (Figure 2-5b). Malgré une floraison abondante, la production de graines est souvent limitée au champ (Yen 1974).

La patate douce est une espèce allogame obligatoire auto-incompatible. L'auto-incompatibilité est de type sporophytique, c'est-à-dire qu'elle s'exprime entre les cellules nourricières de la paroi du grain de pollen (hexaploïde) et celle du stigmate de la fleur (hexaploïde). Elle est attribuée à l'action

Figure 2-4 : Les deux scénarios proposés pour expliquer l'origine du génome hexaploïde de la patate douce (Kobayashi 1984 et Austin 1987).



Hypothèse de Kobayashi: origine autopolyploïde  
à partir du génome de *I. trifida* 2x



Hypothèse de Austin: origine allo ou auto-allo polyplôïde  
à partir de *I. trifida* et *I. triloba*

d'un locus unique S, multi-allélique (Kowyama et al. 1980). Si l'un des allèles de la plante qui a généré le pollen est le même que l'un des allèles du stigmate de la plante receveuse, le développement du tube pollinique est bloqué et la fécondation n'a pas lieu. Seule la fécondation entre des clones appartenant à des groupes d'auto-incompatibilité différents est possible, limitant ainsi les possibilités de croisements réussis au champ. Dix-sept groupes d'auto-incompatibilité ont été identifiés sur un échantillonnage de 700 plantes d'Amérique tropicale et du Pacifique (Nakanishi et Kobayashi 1979). De plus, l'état hexaploïde de la plante induit différents processus de stérilité à différents stades (dérèglements au cours de la méiose par exemple) (Lebot 2009). Pourtant, Yen (1974) atteste que dans certaines régions (notamment aux Philippines et en Nouvelle-Guinée), les cultivateurs reconnaissent les plantules issues de graines et les incorporent fréquemment à leurs stocks de variétés (Figure 2-5c). La forte hétérozygotie des clones de patate douce, en relation avec l'état polyploïde et le croisement obligatoire de clones distincts, permettrait d'obtenir un large éventail de recombinaisons génétiques et de variants morphologiques sélectionnables par les agriculteurs (Yen 1974).

En outre, c'est une espèce particulièrement sujette à la mutation somatique et ce pour ses différents organes (feuilles, tiges et racines) en relation notamment avec son haut niveau de polyploïdie (Lewthwaite 2004). Par exemple, le taux de mutation de la couleur de la racine varie entre 1 et 18% en fonction du cultivar alors que chez la pomme de terre il varie seulement entre 0.0005 et 0.001 %. Ces mutations somatiques sont une source non négligeable de variabilité directement sélectionnable par les agriculteurs. Il a été montré par exemple que les variétés du groupe *Waina*, variétés commerciales majoritaires en Nouvelle-Zélande, dériveraient par mutations somatiques d'un clone introduit en 1860 par les baleiniers (Lewthwaite 2004).

Deux sources de variabilité sont donc susceptibles d'agir au champ et ont joué des rôles différents selon les contextes agro-écologiques (Yen 1974).

## II. DIFFUSION ET DIVERSIFICATION DE LA PATATE DOUCE EN OCEANIE

L'histoire de la diffusion de la patate douce en Océanie est fascinante à plus d'un titre. D'abord d'un point de vue historique et migratoire puisque sa présence en Polynésie avant l'arrivée des Européens reste la preuve la plus convaincante de l'existence de contacts préhistoriques entre les Polynésiens et l'Amérique (voir l'introduction générale). Aussi, la patate douce semble avoir joué un rôle primordial dans l'adaptation des sociétés océaniques aux environnements multiples des îles qu'ils ont colonisées, complétant les possibilités adaptatives des autres plantes introduites depuis l'ouest (le taro et l'igname notamment) (Yen 1974). C'est en Nouvelle-Guinée qu'on trouve aujourd'hui la plus grande diversité de formes de patates douces, faisant de l'île un centre de diversité secondaire pour cette espèce (Bourke 2005, 2009).

### II.1. L'histoire des introductions de la patate douce en Océanie : L'hypothèse tripartite, un résumé

En 1957, Jacques Barrau réalise une revue des preuves archéologiques, linguistiques et historiques sur l'introduction de la patate douce en Océanie et pose les bases de l'hypothèse dite tripartite, puisque trois « routes » principales sont susceptibles d'avoir introduit la patate douce en Océanie. Yen dans sa célèbre monographie « *The Sweet Potato in Oceania* » (1974) poursuit le travail entamé et propose une réévaluation complète de cette histoire en s'appuyant en outre sur une étude détaillée des patrons de variations morphologiques des cultivars du Pacifique. Enfin en 2005, Green revisite cette hypothèse à la lueur des dernières données. Des introductions à la fois préhistoriques et historiques doivent être envisagées pour expliquer les patrons de distribution de la patate douce dans le Pacifique (Figure 2-6).

### Figure 2-5 : Eléments de reproduction sexuée de la patate douce.

Bien que multipliée clonalement par les cultivateurs, la patate douce a conservé la possibilité de se reproduire par voie sexuée.

a) L'appareil reproducteur est composé de fleurs campanulées, de couleur violette, blanche ou rouge, axillées par deux bractéoles étroites. Le pédoncule mesure 3 à 15 cm de long. Le calice est formé de 5 sépales soudés et terminés en pointe. La corolle est en forme d'entonnoir d'environ 2-5 cm de diamètre et 3-6 cm de hauteur, et composée de 5 pétales soudés. Les 5 étamines sont de taille inégale, et comportent des anthères étroitement sagittées et portées par des filets de 4 à 8 mm de long. L'ovaire, long d'environ 1,5 mm, est ovoïde et constitué de deux loges ovariennes biovulées, volumineuses et séparées par une mince cloison. Il est surmonté d'un style filiforme, long de 12 mm, qui se termine par un stigmate bilobé papilleux. Les fleurs s'ouvrent le matin très tôt et restent ouvertes pour quelques heures, puis se ferment et fanent avant midi. La pollinisation est réalisée par des insectes, en particulier par les abeilles.



b) La fécondation aboutit à la formation de fruits qui sont des capsules pouvant contenir jusqu'à 4 graines mais souvent, seulement une ou deux sont bien développées. Malgré une floraison relativement abondante, la production de graines est limitée au champ, en raison i) du système d'auto-incompatibilité sporophytique ii) de problèmes de stérilité et de dérèglements méiotiques liés au niveau de ploïdie. Les graines sont noires, glabres, plates sur un côté et convexes sur l'autre. Le tégument est très dur et la germination est alors très difficile. La germination peut être améliorée par scarification mécanique ou chimique (acide sulfurique).



### II.1.1. Les introductions préhistoriques ou la route Kumara

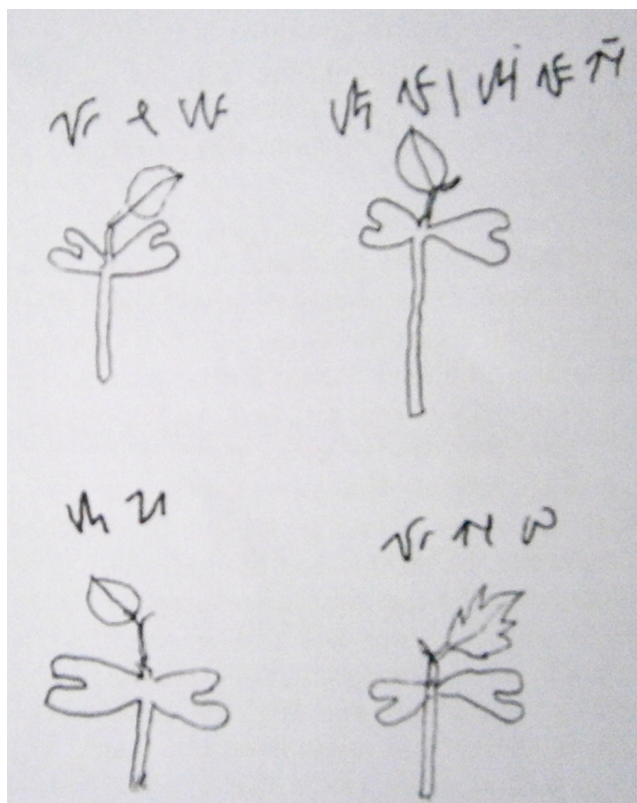
Très tôt, des travaux de linguistique montrèrent que le terme Quechua « *cumar* » (et ses variants) qui désigne la patate douce dans les hautes terres d'Equateur, du Pérou et même de Colombie (Yen 1974), était similaire à celui utilisé dans toute la Polynésie (« *kumara* » et ses variants). Cependant, l'absence de ce nom dans les basses terres des côtes d'Amérique du Sud donnait au départ peu de poids à l'argumentaire linguistique (Brand 1971). Ce terme aurait en effet tout aussi bien pu être transmis plus tard par les voyageurs Européens qui l'auraient diffusé avec la plante. Scaglione (2005) a récemment levé le voile sur cet apparent paradoxe et montré que les populations Canaris d'Equateur, qui s'étendaient à l'époque des contacts présumés (1000 après JC) sur les côtes du Golf de Guayaquil, nommaient la patate douce *comal* ou *cumal*. En outre, des restes de patate douce ont été retrouvés dans différents sites archéologiques du Pacifique, qui pré-datent l'arrivée des Européens, réintroduisant ainsi l'idée d'une diffusion préhistorique. Les plus vieux restes (trois fragments de parenchyme de tubercules carbonisés) proviennent de l'île de Mangaia dans l'archipel des îles Cook. Le plus ancien, daté par contexte par la méthode du radiocarbone ( $^{14}\text{C}$ ), remonte à 988-1155 après JC (Hather et Kirch 1991). D'autres restes archéobotaniques ont ensuite été découverts dans plusieurs sites archéologiques à Hawaï (Coil et Kirch 2005 ; Ladefoged 2005 ; Horrocks et Rechtmann 2009), en Nouvelle-Zélande (Horrocks 2004) et sur l'Île de Pâques (Wallin et al. 2005), suggérant non seulement la présence ancienne de la patate douce en Polynésie mais surtout sa distribution dans tout le triangle polynésien. Les récits des premiers naturalistes européens dans le Pacifique attestent eux aussi de la présence de la patate douce en Polynésie (Yen 1974 ; Green 2005). Elle est identifiée sur l'Île de Pâques en 1722 par Roggeveen. Puis, Joseph Banks et Daniel Solander, les deux botanistes qui accompagnent le premier voyage de James Cook en 1769, et ensuite les Forster, père et fils qui accompagnent les deux expéditions suivantes, font état de sa présence à Tahiti (1769), en Nouvelle-Zélande (1769) et enfin à Hawaï (1778) (Yen 1974). A Tahiti, ils rencontrent notamment une variété jaune qu'ils n'ont à priori jamais vue auparavant. Ils l'identifient comme une nouvelle espèce, *Convolvulus chrysorhizus*, alors qu'à cette époque les autres formes sont référencées sous le nom de *Convolvulus batatas*. Cette variété a été dessinée par S. Parkinson (Figure 2-7a), collectée et sa représentation est aujourd'hui disponible dans l'herbier historique de Banks et Solander, conservé au Muséum d'Histoire Naturelle de Londres (Figure 2-7b). En Nouvelle-Zélande, ils collectèrent aussi une variété à chair blanche (*pallida* pour « pâle » en latin) aux racines allongées, conservée elle aussi dans l'herbier de Londres (Figure 2-7c).

A la lumière des différentes preuves linguistiques, archéologiques et ethnobotaniques, le scénario suivant (**route Kumara**) a été proposé (Green 2005) (Figure 2-6) :

- la patate douce aurait initialement été introduite par des voyageurs polynésiens autour de 1000-1100 après JC en Polynésie orientale, voyageurs qui seraient partis de la région de Mangareva-Pitcairn-Île de Pâques et l'auraient récupérée en Amérique du Sud dans la région du Pérou/Equateur.
- La zone exacte d'introduction initiale est inconnue mais on suppose qu'elle a d'abord été établie quelque part entre les îles Marquises et les îles Cook (dans une région que Yen appelle l'« ellipse centrale »).
- Par la suite, les Polynésiens l'auraient dispersé à Hawaï *via* les îles Marquises entre 1100 et 1300 après JC, puis sur l'Île de Pâques depuis les Touamotou, Mangareva et Pitcairn, à peu près à la même période. Enfin, la plante aurait été introduite en Nouvelle-Zélande durant la phase initiale de colonisation polynésienne, autour de 1200 après JC (Wilmschurst et al. 2011).
- La patate douce aurait aussi pu entrer à Hawaï une seconde fois à la suite d'autres contacts préhistoriques avec l'Amérique, voyages qui auraient atteint les côtes colombiennes (Rensch 1991) ou encore mésoaméricaines. Une variété produisant de grosses racines et à chair blanche, présente seulement à Hawaï au temps des premiers contacts avec les Européens, aurait ainsi pu être introduite.
- Enfin, les Polynésiens auraient pu poursuivre l'introduction de la patate douce plus à l'ouest, en Mélanésie. En Mélanésie orientale (îles Fidji, Vanouatou et Nouvelle-Calédonie), la patate douce est appelée « kumala ». Néanmoins, aucune preuve archéologique ne vient supporter cette hypothèse et les données historiques sont ambiguës (Allen 2005). Aussi, certains linguistes et archéologues défendent l'idée que la « kumara » polynésienne a pu être introduite

c) Une fois les graines tombées, elles peuvent se conserver longtemps dans la banque de graines du sol. Lors de l'ouverture d'un jardin, les graines de patate douce vont germer et de jeunes plantules apparaître. Dans certaines régions, les cultivateurs savent reconnaître ces plantes issues de graines et les incorporent fréquemment à leurs stocks de variétés.

Ci-dessous, des dessins *Hanunó'o* (province de Mindoro aux Philippines) faits par les cultivateurs et représentant les jeunes plantules qu'ils trouvent juste après ouverture de leurs nouveaux jardins. Les deux cotylédons sont bien figurés ainsi que la première vraie feuille, différente selon des variétés.



D'après Yen 1974.



très tôt (il y a environ 1200 ans) en Nouvelle-Guinée (Golson 1977; Gorecki 1986; Haberle 1998; Haberle et Atkin 2005). Le témoignage palynologique du site archéologique de Kuk (dans la vallée de Wahgi située dans les hautes-terres de Nouvelle-Guinée) démontre la disparition progressive de la forêt naturelle il y a 1 200 ans, un déboisement qui peut être interprété comme la conséquence de l'introduction de la patate douce. La sédimentation du sol à cette même période marquerait elle aussi l'arrêt du drainage en relation avec la production de patates douces (Golson 1977). Néanmoins, aucun reste direct de patate douce n'a été identifié à cette période. A l'heure actuelle, cette hypothèse manque elle aussi de preuves tangibles et l'introduction de la patate douce en Nouvelle-Guinée est probablement plus récente, il y a au plus 400-450 ans (Ballard 2005).

Selon Green (2005), au moins trois variétés (lignées) différentes auraient ainsi été introduites par les Polynésiens :

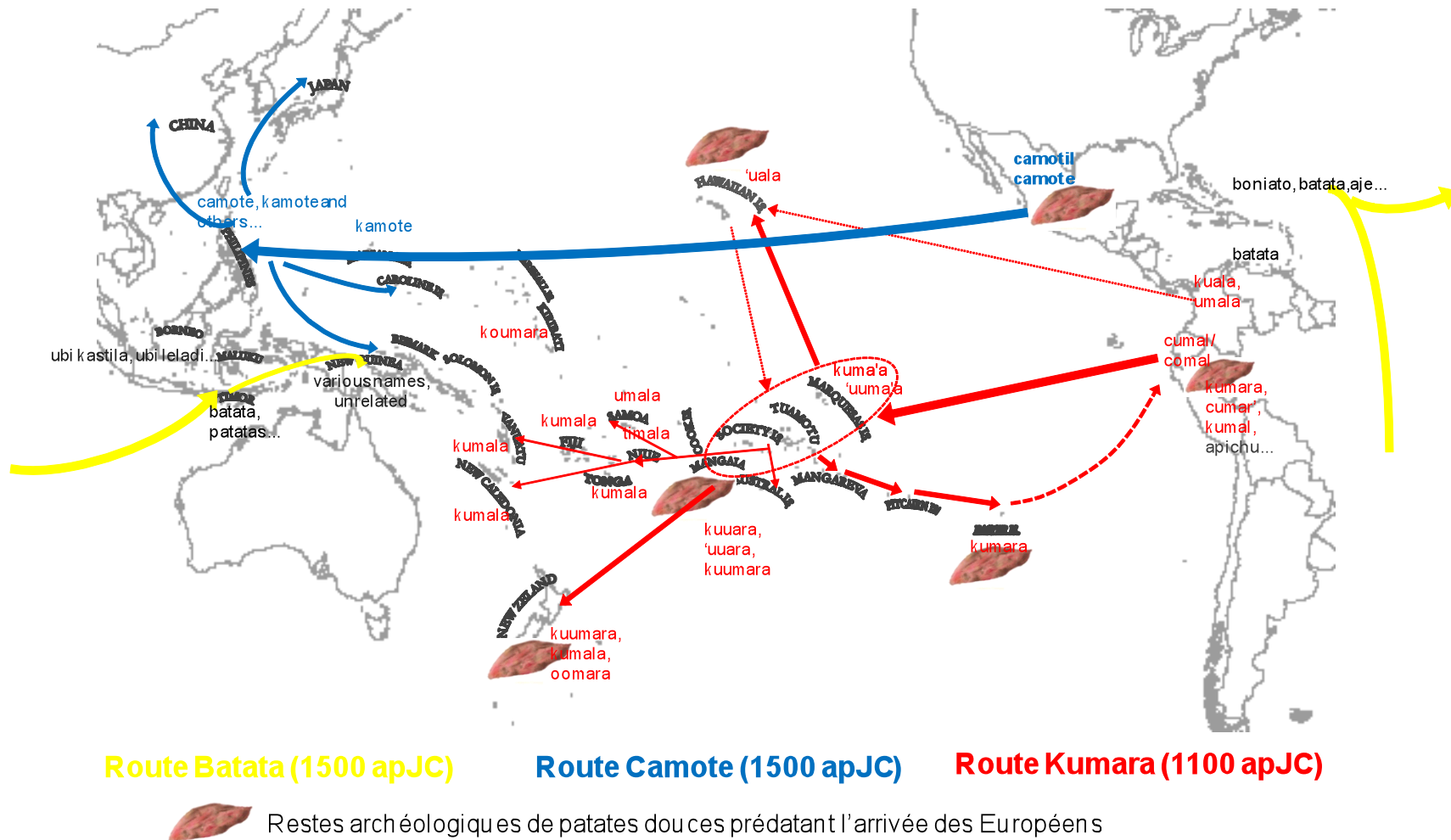
- une variété à chair jaune et une à chair blanche et aux racines allongées qui auraient toutes deux une origine sud-américaine (côtes du nord du Pérou ou d'Equateur)
- une variété à chair blanche avec une forme globuleuse dont l'origine est inconnue (nord de la Colombie, Amérique du Sud, Amérique Centrale ?).

### *II.1.2. Les introductions historiques : les routes Camote et Batata*

- Les galions espagnols faisant route entre Acapulco à Manille dès le 16<sup>ième</sup> siècle auraient ainsi introduit la patate douce aux Philippines depuis le Mexique (**route Camote**). Il est possible que ces mêmes galions aient introduit la patate douce à Hawaii durant ces voyages, certains navires ayant pu s'échouer sur ces îles.
- A la même époque, les Portugais auraient introduit la patate douce en Indonésie depuis leurs colonies caribéennes (**route Batata**)
- Elle aurait alors gagné les côtes de Nouvelle-Guinée dès le 17<sup>ième</sup> siècle suite à des échanges locaux, puis les hautes terres à partir de 1700 AD (Ballard 2005). Dans les basses terres, ainsi que dans les autres îles d'Océanie proche, il semble qu'elle ait été introduite (ou réintroduite) plus tardivement, courant 19<sup>ième</sup> siècle (Bourke 2005).
- Aussi, il est possible que les voyages de Queiros et Mendaña qui débutaient de *Calao* au Pérou, aient introduit la patate douce très tôt dans le Pacifique au 16<sup>ième</sup> siècle (au Vanouatou et/ou peut-être aux Marquises) (Baert 1995 et 2001).
- Dès le 18<sup>ième</sup> et notamment durant les 19<sup>ième</sup> et 20<sup>ième</sup> siècles, la patate douce aurait été fréquemment réintroduite puis redistribuée par les missionnaires, les baleiniers, les commerçants et les insulaires eux-mêmes dans toute l'Océanie (Yen 1974). La direction et l'intensité de ces échanges sont toutefois difficiles à établir et à quantifier précisément.

Ainsi selon cette hypothèse tripartite, les variétés de l'ouest du Pacifique devraient avoir une origine principalement mésoaméricaine et/ou caribéenne, alors que les variétés polynésiennes dériveraient plutôt des cultivars d'Amérique du Sud. Nous avons vu précédemment que l'étude de la structure de la diversité génétique des cultivars d'Amérique tropicale suggère l'existence de deux groupes génétiques différenciés, le groupe mésoaméricain (le plus divers) et le groupe péruvo-équatorien (Zhang et al. 2000 ; Gichuki et al. 2003). Ainsi, il doit être possible de retracer les différentes voies d'introduction de la patate douce dans le Pacifique, voies qui ont introduit des variétés d'origines différentes. Toutefois, il est probable que les transferts historiques aient largement modifié et brouillé ces patrons initiaux de distribution. Les nouvelles introductions ont pu remplacer les anciennes variétés (Barrau 1957 ; Yen 1974), compliquant dès lors la reconstruction phylogéographique de l'histoire de l'introduction de la patate douce dans le Pacifique.

Les termes utilisés pour nommer la patate douce dans les différentes régions du Pacifique sont présentés ainsi que la localisation des restes archéologiques directs retrouvés dans le triangle Polynésien.





## II.2. Diversités océaniques ; une histoire d'adoption, d'adaptation, et de diversification

Aujourd'hui, la patate douce est abondamment distribuée partout dans le Pacifique (à l'exception des Atolls des Touamotou) et ne semble donc pas présenter d'exigences écologiques particulières, contrairement à d'autres plantes comme le taro et l'igname. Le taro est en effet à la base des agrosystèmes des zones humides et l'igname de ceux des zones plus sèches (Barrau 1965). La patate douce possède une large adaptabilité agronomique, à la fois aux conditions climatiques et édaphiques. Et c'est précisément cette caractéristique intrinsèque, plutôt que l'ancienneté de son introduction, qui semble avoir conditionné les patrons d'adoption de cette plante en Océanie (Yen 1974). Elle aurait ainsi principalement prospéré dans les régions marginales aux conditions défavorables à la culture des autres plantes de subsistance déjà présentes, prospérant ainsi dans les zones particulièrement arides et les régions montagneuses.

À Hawaii, les premiers explorateurs décrivent la patate douce comme une culture particulièrement abondante : « *it is in such plenty that the poorest natives would throw them into our ships for Nothing* » (le Capitain King en 1779, cité dans Yen (1974)). L'analyse des sites agricoles archéologiques hawaïens (notamment les sites de Kohala et Kula sur l'île d'Hawaii ou encore Kahikinui sur l'île de Maui) atteste dès 1 400 après JC qu'elle y est devenue la culture de base des systèmes agricoles des versants secs, sous le vent, des jeunes îles volcaniques, et ce particulièrement dans les hautes terres. Par contre, elle est absente des sites côtiers sous le vent où prospèrent les parcelles irriguées de taro (Kirch et Coil 2005 ; Ladefoged 2005). En Nouvelle-Zélande, la patate douce constitue la culture principale des Maoris à l'arrivée des Européens. Elle aurait été introduite en Nouvelle-Zélande avec les premiers migrants et leur aurait permis de s'installer sur cette île au climat tempéré, très peu propice à la culture des autres plantes tropicales. Sur l'Île de Pâques, une île certes tropicale mais au climat rigoureux et incertain, l'introduction de la patate douce dès 1200 après JC a permis l'utilisation de terres jusque-là non productives, les terres de l'intérieur de l'île tout comme les hautes terres, et entraîné une production de nourriture en surplus (Wallin et al. 2005). De même, l'adoption de la patate douce en Nouvelle-Guinée a permis aux populations locales de s'étendre plus en altitude dans les hautes terres (Ballard 2005). En revanche, à l'arrivée de Cook à Tahiti, la patate douce semble être une plante de seconde importance (Green 2005). Cette même description est faite plus tard, au début du 20<sup>ème</sup> siècle par Henry (cité dans Brown 1935) et relève l'existence de seulement quelques variétés peu adaptées aux conditions humides de l'île. Cette situation semble d'ailleurs être valable pour la plupart des autres îles de Polynésie orientale, comme les Samoa ou les Marquises (Brown 1935). De même en Mélanésie elle semble occuper une place secondaire. Dans ces régions où la culture du taro et de l'igname prospère, ces plantes anciennes demeurent les espèces valorisées culturellement.

D'un point de vue agronomique, la patate douce ressemble aux autres plantes à racines et tubercules déjà cultivées sur les îles du Pacifique, et a pu être intégrée relativement facilement aux systèmes agricoles en place lors de son introduction (Yen 1974). Malgré tout, elle n'est pas une plante de région froide et le développement de pratiques culturelles appropriées a été nécessaire à son adaptation dans certaines régions. Ainsi en Nouvelle-Zélande, les Maoris solutionnent le problème du froid en cultivant sur l'île du Nord (l'île du Sud restant trop froide) la patate douce comme une plante annuelle, en paillant les buttes de culture et conservant les racines dans des fosses creusées dans le sol pour leur consommation hivernale et pour leurs futures boutures (Yen 1974) (Figure 2-8a). Aussi dans les hautes terres de Nouvelle-Guinée, la culture de la patate douce est techniquement élaborée. Elle y est réalisée en aménageant le sol en billons (ou en buttes), toujours entourés de larges fossés de drainage car elle est surtout cultivée dans les bas-fonds des vallées. Paille et boue organique sont alors régulièrement apportées et constituent un véritable compostage qui réchauffe le sol (Barrau 1962) (Figure 2-8b).

En retour, la performance de cette nouvelle agriculture vivrière dans ces différentes régions jusque-là peu ou pas exploitées a conduit à d'importantes transformations et innovations socioculturelles (Ballard 2005). En Nouvelle-Guinée, l'introduction de la patate douce est rapidement suivie d'une forte croissance démographique et de nombreux mouvements migratoires. Dans un second temps, de larges transformations culturelles et sociales se mettent en place, certains auteurs

**Figure 2-7 : Spécimens de patates douces observés lors du premier voyage de J. Cook dans le Pacifique en 1769.**

- a) *Convolvulus chrysorhizus*, une variété de patate douce à chair jaune, répandue à l'époque dans tout le triangle polynésien. Dessin réalisé par S. Parkinson à Tahiti entre le 13 avril et le 9 Aout 1769.



© Natural History Museum, London.

parlant même de « révolution ipomoéenne » (Ballard 2005). On y observe notamment une redéfinition des interactions sociales, des réseaux d'échanges et des cérémonies associées, où le porc joue un rôle prépondérant (Wiessner 2005, Brown et Brookfield 2005). La patate douce sert en effet à nourrir les hommes mais aussi les porcs et son introduction aurait d'ailleurs permis d'en intensifier l'élevage. L'élevage du porc occupait les hommes, la culture des patates douces les femmes, participant ainsi à une redéfinition de la distribution des pratiques agricoles (Brown et Brookfield 2005). Sur l'Île de Pâques, cette production de nourriture en surplus aurait notamment permis l'avènement des sculptures monumentales « *moai* » et plates-formes lithiques « *ahu* » caractéristiques de l'île (Wallin et al. 2005).

L'adaptation de la patate douce en Océanie n'est pas seulement déterminée par des facteurs environnementaux et culturels, mais aussi biologiques et génétiques, qui contraignent les possibilités intrinsèques de diversification de la plante et donc la sélection par les agriculteurs de nouveaux cultivars adaptés. Nous avons vu précédemment que deux sources de diversification sont envisageables, la mutation somatique ou la recombinaison génétique. A Tahiti, au début du 20<sup>ème</sup> siècle, seulement six variétés sont décrites et à Rarotonga (îles Cook) seulement deux (Brown 1935). Plus tard dans les années 1960, Yen collecte une centaine de variétés en Polynésie orientale mais la plupart, bien qu'ayant des noms différents, sont très probablement des lignées clonales (Yen 1974). Par contre à Hawaï au début du 20<sup>ème</sup>, Handy (1940) dresse une longue liste de plus de 170 variétés de patates douces dont 25 au moins seraient des variétés anciennes, les nombreuses autres étant soit des sélections locales de mutants, soit des plantes issues de graines, mais le plus souvent des doublons renommés après déplacements, et aussi des variétés introduites plus tard par les Européens. A la même période, Best (1921) recense presque 100 noms de variétés en Nouvelle-Zélande. En Nouvelle-Guinée, plus de 5 000 variétés pourraient être aujourd'hui cultivées. Il semble donc que la distribution de la diversité des variétés de patate douce océaniques soit relativement hétérogène. Cette situation reflète probablement des trajectoires de diversification variées où diversification clonale et diversification sexuée ont joué des rôles différents. En Polynésie, peu de clones ont probablement été introduits depuis l'Amérique, limitant la base génétique initiale. Dès lors, les possibilités de recombinaison et de sélection de nouveaux variants pour cette espèce auto-incompatible devaient être limitées. Comment expliquer alors la situation rencontrée à Hawaï, ou encore en Nouvelle-Zélande, où la patate douce semble très tôt relativement bien diversifiée ? D'autres introductions polynésiennes auraient-elles élargi la base génétique de départ ou bien la sélection attentionnée de mutants somatiques aurait-elle permis l'obtention de ces nombreuses variétés ? A l'ouest, les introductions historiques ont probablement amené un matériel de départ plus diversifié. Ces introductions ne semblent pourtant pas suffisantes pour expliquer l'incroyable diversité rencontrée aujourd'hui en Nouvelle-Guinée et aux Philippines (Yen 1974). Il est probable que dans ces régions, les nombreuses variétés recensées chez les cultivateurs doivent leur existence à des semis accidentels plutôt qu'à des mutations somatiques ou de nouvelles introductions.

### II.3. A la recherche de preuves génétiques

L'histoire de la patate douce en Océanie a été principalement retracée au travers d'arguments linguistiques, archéologiques, anthropologiques et ethnobotaniques (revue récemment par Ballard et al. 2005). En revanche, peu d'études génétiques se sont attelées à une reconstruction phylogéographique de sa diffusion (Zhang et al. 1998 ; Fajardo et al. 2002 ; Zhang et al. 2004 ; Clarke 2009). Selon l'hypothèse tripartite, on s'attend notamment à différencier génétiquement les variétés de l'est du Pacifique, proches de celles de la région Pérou-Equateur, de celles de l'ouest du Pacifique plutôt dérivées des variétés d'Amérique Centrale et des Caraïbes. Aussi, on peut s'attendre à une diversité réduite et une évolution majoritairement clonale en Polynésie. En revanche, en Mélanésie la base génétique est probablement plus large et la recombinaison génétique aussi.

Zhang et al. (2004) ont génotypé 76 cultivars dont 31 cultivars d'Océanie, 12 des Philippines, 22 du Pérou-Equateur et 11 du Mexique à l'aide de marqueurs AFLP. Ils montrent que les cultivars océaniques sont proches génétiquement des variétés mexicaines mais distants de celles provenant de la région Pérou-Equateur. Selon ces auteurs, ces résultats sont en faveur d'une origine essentiellement mésoaméricaine des cultivars océaniques. Toutefois, leur échantillonnage est biaisé car il intègre

- b) Spécimens d'herbier de *Convolvulus chrysorhizus* récoltées à Tahiti en 1769 par J. Banks et D. Solander, les deux botanistes qui accompagnent le premier voyage de James Cook. Photographie prise au Muséum d'Histoire Naturelle de Londres, herbier historique de J. Banks.



principalement des variétés de Papouasie-Nouvelle-Guinée et des îles Salomon, quelques variétés de Tonga (7), et seulement trois de Polynésie orientale (îles Cook).

Récemment, Clarke (2009, thèse publiée publiquement en 2012) a génotypé 246 accessions provenant de la collection de Douglas Yen avec des marqueurs AFLP, intégrant cette fois-ci une quarantaine de variétés polynésiennes. L'arbre de diversité obtenu révèle une structuration géographique des cultivars relativement forte. Les variétés polynésiennes constituent deux lignées bien différenciées des autres variétés de l'ouest du Pacifique (Mélanésie, Nouvelle-Guinée), relativement peu différenciées entre elles et proches génétiquement. Ces résultats suggèrent que les variétés polynésiennes ont une origine distincte de celle de l'ouest du Pacifique et sont interprétés par l'auteur comme la marque des introductions préhistoriques des *kumaras* polynésiennes depuis l'Amérique du Sud. Toutefois, en l'absence d'une représentation exhaustive des régions américaines d'où dérivent les variétés océaniques (la région Pérou-Equateur pour les variétés polynésiennes, le Mexique et les Caraïbes pour les variétés de l'Ouest du Pacifique), il demeure prématuré de statuer sur leur origine.

Quelques études se sont intéressées à la caractérisation de la diversité génétique en Papouasie-Nouvelle-Guinée. Zhang et al. (1998) montrent que les accessions de Papouasie sont clairement différenciées des variétés sud-américaines et ont un niveau de diversité inférieur. Fajardo et al. (2002) étudient la diversité de 141 plantules de patates douces issues de graines ramassées dans les champs des agriculteurs des hautes terres de Nouvelle-Guinée. La base génétique apparaît relativement étroite et non structurée géographiquement, à l'exception de quelques plantes de l'archipel Bismarck.

Malgré ces différentes tentatives, il manque à l'heure actuelle une étude phylogéographique exhaustive de la patate douce en Océanie qui permette à la fois de tester l'hypothèse tripartite mais aussi d'éclairer plus finement les processus de diversification qui ont façonné les patrons de diversité océaniques. Cette caractérisation des ressources génétiques apparaît d'autant plus essentielle, considérant l'importance grandissante de la patate douce dans les systèmes vivriers océaniques. Le glissement de l'agriculture de subsistance vers une agriculture commerciale a mis en relief la supériorité de la patate douce sur les autres cultures traditionnelles. En outre, là où la population augmente rapidement, la préférence accordée à la patate douce témoigne de sa capacité de production. Aussi, les changements de mode de vie de ces sociétés s'accompagnent souvent d'une réduction du temps alloué à la production vivrière et les insulaires se tournent donc vers des cultures qui demandent moins de travail et qui ont des cycles végétatifs plus courts (quatre mois pour la patate douce contre huit à neuf mois pour les ignames, taro, macabo et manioc).

### III. PROBLEMATIQUE(S) ET PLAN DE THESE

L'objectif général de cette thèse est de retracer l'histoire évolutive de la patate douce depuis sa domestication en Amérique tropicale jusqu'à son introduction et sa diversification en Océanie. C'est par le prisme de la génétique que nous abordons ici cette histoire, au travers de l'étude de la distribution de la diversité génétique neutre de marqueurs chloroplastiques (microsatellites et séquences intergéniques) et nucléaires (microsatellites), marqueurs qui permettent d'avoir accès à des histoires évolutives distinctes et complémentaires. Nous menons ici une enquête phylogéographique aux origines de la diversité de *Ipomoea batatas*.

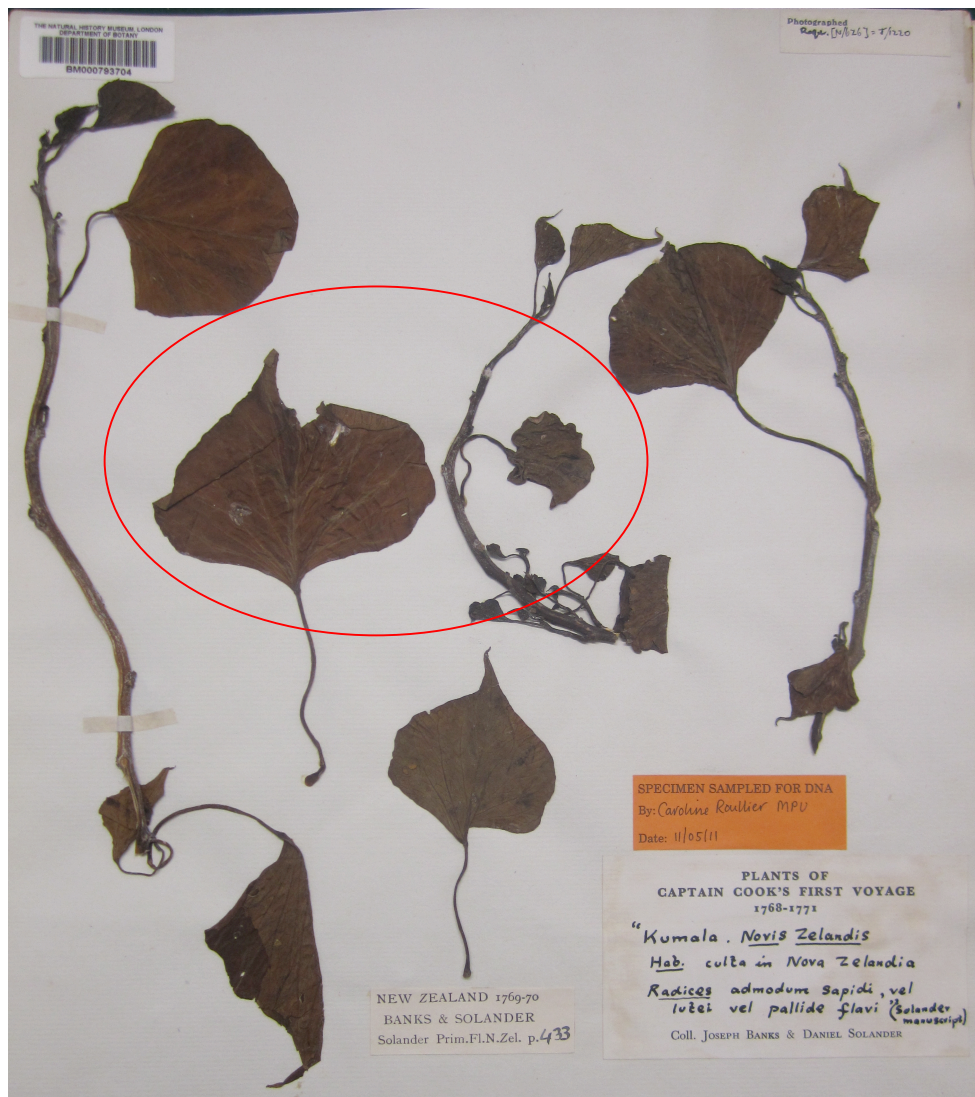
Cette enquête s'articule autour de deux grands axes, chacun composé de deux chapitres. En tout, quatre chapitres composés chacun d'un manuscrit (en anglais car correspondant à autant d'articles) constituent le corps de ce travail.

#### - l'histoire de la domestication de la patate douce en Amérique tropicale :

La première étape de cette enquête (**chapitre 3**) consiste à caractériser la diversité génétique des cultivars d'Amérique tropicale (du Mexique au Pérou) et à appréhender la question de l'origine géographique de la patate douce. Les données génétiques disponibles, toujours obtenues pour un nombre restreint d'échantillons, suggèrent l'existence de deux groupes distincts, l'un correspondant



c) Spécimen de patate douce collectée en Nouvelle Zélande par J. Banks et D. Solander en 1769. C'est une variété de kumara à chair blanche (*pallida* pour « pâle » en latin) aux racines allongées. Photographie prise au Muséum d'Histoire Naturelle de Londres, collection permanente.



aux variétés de la région Pérou-Equateur et l'autre à celle d'Amérique Centrale et des Caraïbes. Cette différenciation est généralement interprétée comme le résultat de la diffusion de la patate douce depuis un centre de domestication unique situé quelque part entre le Mexique et le Venezuela et d'une diversification secondaire en Amérique du Sud. Pourtant, les plus vieux restes archéologiques sont enregistrés dans l'ouest de l'Amérique du Sud sur la côte péruvienne. Nous cherchons donc dans ce chapitre à savoir si la structure actuelle de la patate douce en Amérique tropicale provient de foyers indépendants de domestication ou a une origine unique. En outre, cette caractérisation génétique est essentielle pour « pister » par la suite, l'origine des variétés océaniques et tester l'hypothèse tripartite.

Les patrons de diversité obtenus dans le chapitre 3 suggèrent fortement une origine multiple en Amérique Centrale et en Amérique du Sud. Toutefois, la patate douce est un taxon hexaploïde pour lequel les différents contributeurs sauvages ne sont pas clairement identifiés. Deux hypothèses sont couramment envisagées, une origine autopolyploïde à partir d'un taxon sauvage diploïde *I. trifida*, et une origine allo-polyploïde ayant impliqué l'hybridation de deux espèces distinctes (*I. trifida* et *I. triloba*). Des populations sauvages tétraploïdes de *I. batatas* ont parfois été décrites par les taxonomistes, mais très souvent négligées dans les études génétiques, alors qu'elles représentent probablement un maillon essentiel pour la compréhension de l'origine botanique de la patate douce. Le signal polyphylétique détecté dans le chapitre 1 pourrait être i) le résultat de domestications indépendantes (de populations différenciées conspécifiques ou d'espèces distinctes), ii) le résultat de phénomènes d'hybridation post-domestication. Aussi, le rôle de la polyploïdisation dans le processus de domestication est encore peu documenté. Pour certaines espèces, il est clair que la polyploïdisation a précédé de loin la domestication. Pour d'autres, les mouvements et transferts de plantes par les hommes, au cours d'une phase dite de pré-domestication, ont permis de mettre en contact des espèces distinctes, à l'origine de phénomènes d'hybridation inter-spécifique spontanée (cf. introduction générale). **Dans le chapitre 4**, nous cherchons donc à clarifier l'origine botanique de la patate douce ainsi que les relations entre domestication et polyploïdisation.

A l'issu de ces deux premiers chapitres, la distribution de la diversité en Amérique tropicale est bien caractérisée et nous pouvons nous lancer dans l'étude de l'histoire de la diffusion et de la diversification de la patate douce en Océanie.

## - l'histoire de la diffusion et de la diversification de la patate douce en Océanie

L'enquête se poursuit alors en Nouvelle-Guinée, centre de diversité secondaire de la patate douce où plus de 5 000 cultivars différents seraient aujourd'hui cultivés. La Nouvelle-Guinée se situe potentiellement à la rencontre des trois voies d'introduction de la plante en Océanie (Kumara, Batata, Camote) même si les preuves d'une introduction préhistorique sont actuellement réduites. Dans cette région, Yen fait l'hypothèse que les nombreuses variétés recensées doivent leur existence à la sélection par les agriculteurs de plantules issues de graines plutôt qu'à l'utilisation de mutations somatiques, ou à des échanges importants entre régions et introductions nouvelles. **Dans le chapitre 5**, nous réalisons une étude de la structure de la diversité génétique de plus de 400 variétés de Nouvelle-Guinée que nous comparons à celle présente en Amérique tropicale pour i) identifier l'origine des variétés initialement introduites en Nouvelle-Guinée, ii) caractériser les processus de diversification en jeu dans ce centre de diversité secondaire et notamment évaluer l'importance de la diversification sexuée et clonale et enfin iii) étudier la distribution géographique de la diversité allélique sur cette île qui présente une forte hétérogénéité environnementale, en vue notamment d'aider à la gestion des collections *ex situ* maintenues en Papouasie-Nouvelle-Guinée.

Finalement, dans le dernier chapitre (**chapitre 6**), nous nous attachons à tester l'hypothèse tripartite d'introduction de la patate douce en Océanie. Dans la lignée d'un certain nombre de travaux récents qui cherchent à reconstruire les migrations des hommes au travers de la phylogéographie de leurs espèces domestiques (cf. introduction générale), nous partons à la recherche de preuves génétiques de contacts transocéaniques anciens entre la Polynésie et l'Amérique et de la route d'introductions préhistoriques Kumara. Nous réalisons une phylogéographie exhaustive des cultivars d'Océanie et d'Amérique tropicale et comparons la distribution de la diversité génétique (chloroplastique et nucléaire) d'un échantillonnage moderne (variétés collectées depuis les années 1960 jusqu'à aujourd'hui et maintenues en collections *ex situ*) à des échantillons d'herbiers collectés

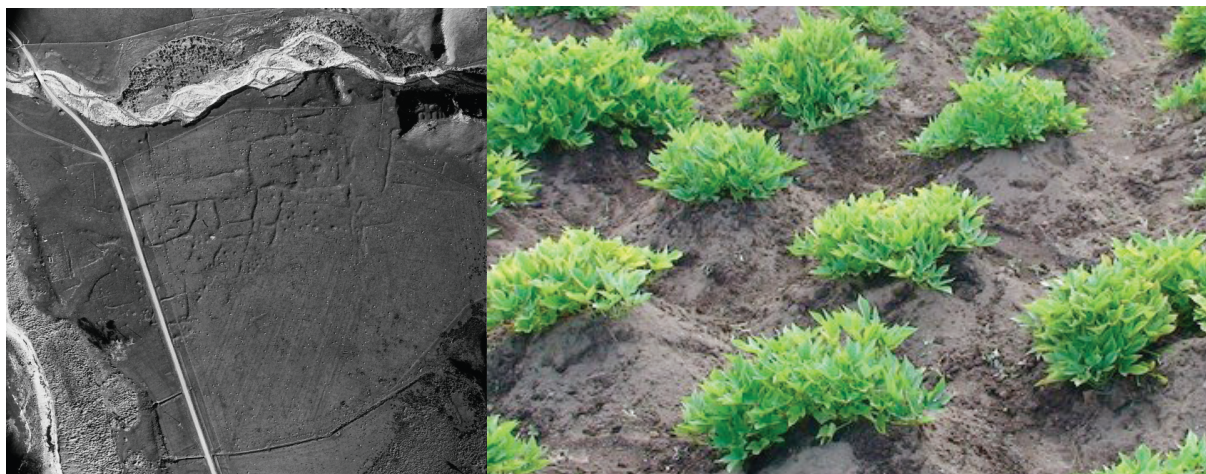


**Figure 2-8 : Jardins de patate douce.**

a) Jardins de patate douce vus du ciel dans les hautes terres de Nouvelle Guinée (région de Moro). Les buttes peuvent atteindre jusqu'à deux mètres de diamètre. Plusieurs variétés sont généralement plantées sur chaque butte.



b) Vue du ciel d'anciens jardins côtiers de patates douces situés au Sud-Est de l'île du Nord de Nouvelle Zélande (panel de gauche). Les jardins étaient délimités par des murs de pierres qui servaient notamment de support aux haies végétales « coupe vent ». Reconstitution d'un jardin Maori (panel de droite). Les buttes mesurent environ 80 cm et un pied de patate douce pousse sur chacune des buttes. Le sol y était longuement travaillé et notamment enrichi en petites pierres et en sable.





en Polynésie dès 1769, dont ceux rapportés par J. Cook lors de son premier voyage à bord de l'*Endeavour*. L'utilisation de ces échantillons d'herbier permet d'avoir accès aux premières variétés introduites sur les îles Polynésiennes et ainsi de s'affranchir de l'effet des introductions historiques. Dans ce dernier chapitre, il est aussi question de mener une analyse régionale des patrons de diversité et de discuter et comparer les processus de diversification en jeu dans ces différentes régions, notamment entre la Polynésie et la Mélanésie.

Dans une discussion générale, nous reconstruirons pas à pas l'histoire évolutive de la patate douce de l'Amérique tropicale à l'Océanie et discuterons les limites de notre approche méthodologique. Il sera aussi question d'évaluer les implications de ce scénario pour la gestion et la conservation des ressources génétiques.

## IV. LES OUTILS DE L'ENQUETE

### IV.1. Les plantes

Au vu de l'étendue de la zone considérée dans ce travail, ainsi que de la difficulté d'accès à certaines régions, notre échantillonnage s'est principalement concentré sur du matériel végétal conservé dans des collections *ex situ*.

La reconnaissance internationale de l'importante alimentaire de la patate douce a conduit à l'élaboration de nombreuses collections *ex situ* de par le monde. Plus de 29 000 accessions seraient conservées dans ces collections (FAO 2011). Le CIP, Centre International de la Pomme de terre situé à Lima au Pérou, joue un rôle primordial dans la conservation et l'amélioration des ressources génétiques de la patate douce. Plus de 5 526 accessions, comprenant notamment de nombreuses variétés locales et provenant de plus de 57 pays, y sont actuellement conservées au champ ainsi que *in vitro*. Plus d'un millier d'accessions de *Ipomoea* sauvages, dont les ancêtres présumés de la patate douce y sont aussi conservées sous forme de graines. Plusieurs autres collections nationales qui conservent des ressources génétiques locales principalement aux champs ont attiré notre attention. Parmi elles, il y a les collections du NARI, l'Institut National de Recherche Agronomique de Papouasie-Nouvelle-Guinée, qui comptent plus de 1 000 variétés traditionnelles collectées à la fois dans les agrosystèmes des basses et des hautes terres. Aussi, Douglas Yen dans les années 1960 a constitué une vaste collection des variétés d'Océanie qui inclut également des cultivars d'Asie du Sud-est (notamment de Philippines et d'Indonésie) ainsi que quelques variétés du nord de l'Amérique du Sud (Colombie, Pérou, Equateur). Cette collection est bien référencée et a fait l'objet d'études morpho-agronomiques et historiques détaillées (Yen 1974). Elle est aujourd'hui maintenue à l'Institut National des Sciences du Végétal de Tsukuba au Japon. En outre, l'institut national des sciences agronomiques (NIAS) de Tsukuba maintient plusieurs accessions de *Ipomoea* sauvages et notamment certains spécimens des formes intermédiaires échappées de culture et/ou sauvages de *I. batatas* (tétraploïdes et hexaploïdes), particulièrement intéressantes pour la reconstruction de l'origine botanique de la patate douce. Aussi, une collection des variétés du Vanouatou a été récemment élaborée par Julie Sardos en 2008 et elle est maintenue au Centre Technique de Recherche Agronomique du Vanouatou (CTRAV) à Santo. Enfin, une collection comprenant une soixantaine de variétés locales des îles de Polynésie française (archipels des îles Marquises, Australes et de la Société) a été constituée en 2011 par le service du développement rural (SDR) de Papa'ara à Tahiti. Au total, nous avons ainsi pu accéder à pratiquement 1200 variétés locales préservées dans ces cinq collections *ex situ*, ainsi qu'à 160 accessions correspondant à cinq taxons sauvages apparentés (*I. trifida*, *I. triloba*, *I. tabascana*, *I. leucantha*, *I. tiliacea*) et à quelques formes échappées de culture et/ou sauvages de *I. batatas*. Une liste détaillée de ces accessions, des collections d'origine et de leurs données passeport complètes est fournie en annexe 1.

Les mouvements européens ont probablement affecté, voire complètement effacé, la plupart des patrons initiaux de distribution des plantes en Océanie. Reconstruire les événements de diffusion anciens avec des approches moléculaires nécessite dès lors d'avoir une approche spatio-temporelle,

c'est-à-dire de recourir à l'utilisation d'échantillons anciens. Tandis que les échantillons archéologiques sont fréquemment utilisés pour la reconstruction de l'histoire évolutive des animaux domestiques comme par exemple le poulet (Storey et al. 2007) ou le porc (Larson et al. 2007, 2010), la chèvre (Fernandez et al. 2006), ou encore le chien (Wayne et al. 2006), mais aussi des plantes à graines comme le maïs (Freitas et al. 2003 ; et Jaenicke-Després et al. 2003), la disponibilité de tels échantillons pour les plantes à racines et tubercules en milieu tropical est souvent réduite. Les restes archéobotaniques sont rares et se résument au mieux à quelques morceaux de parenchyme de racines ou tubercules carbonisés, le plus souvent des grains d'amidon, des phytolithes ou encore des fragments de vaisseaux de xylème peu appropriés à l'extraction d'ADN. En revanche, les herbiers historiques représentent une source intéressante de matériel végétal comme en témoigne une récente étude sur l'histoire controversée de l'introduction de la pomme de terre en Europe (Ames et Spooner 2008). Toutefois, les herbiers historiques sont rarement utilisés, probablement car l'accès à ces herbiers pour des prélèvements « destructifs » de matériel végétal est extrêmement restreint. Dans notre cas, les spécimens d'herbiers collectés par les naturalistes qui ont accompagnés les expéditions exploratoires et scientifiques des 18 et 19<sup>ème</sup> siècles nous permettraient de retracer dans le temps et l'espace l'évolution des patrons de distribution et donc d'évaluer l'impact des mouvements européens. Bien que bon nombre d'entre eux ont été perdus ou sont éparpillés dans de nombreuses collections dans le monde, certaines planches sont bien identifiées et conservées dans les collections historiques des jardins botaniques et musées d'histoire naturelle. Ainsi, l'herbier de J. Banks et D. Solander, réalisé lors du premier voyage du capitaine Cook à bord de l'*Endeavour*, est conservé au Musée d'Histoire Naturelle de Londres. Il comporte les spécimens collectés à Tahiti en 1768 et en Nouvelle-Zélande en 1771. Ces spécimens collectés au temps des premiers contacts permettent d'avoir accès aux variétés introduites par les Polynésiens pré-datant l'arrivée des Européens sur ces îles. De même, une partie de l'herbier de P. Commerson, naturaliste qui accompagna L.A. de Bougainville dans son voyage autour du monde entre 1768-1769, est maintenue au Muséum National d'Histoire Naturelle de Paris ainsi qu'à l'herbier de l'institut botanique de Montpellier. Toutefois, aucune variété océanienne n'a pu y être identifiée. Aussi, celui de l'expédition Malaspina (1789-1794) est disponible au Jardin Botanique de Madrid. Mais là encore, aucune *Ipomoea batatas* n'a été repérée. Plus tard, notamment au début du 20<sup>ème</sup> siècle, plusieurs botanistes, ethnologues ou agronomes, réaliseront eux aussi des herbiers des plantes du Pacifique (FR Fosberg ESC Handy, FBH Brown, ED Merrill, JFG Stockes ...). Leurs planches d'herbiers sont éparpillées dans les collections permanentes de nombreux herbiers dans le monde (notamment l'herbier des Jardins Botaniques Royaux de Kew, l'herbier du Bishop Museum à Hawaii et l'herbier de l'Université de Leiden) (Pays-Bas) Ces échantillons sont tout aussi intéressants car ils permettent de suivre l'évolution de l'influence des transferts de l'aire européenne dans le Pacifique.

Nous avons donc parcouru certains de ces herbiers à la recherche de spécimens océaniens de patate douce. Au total, une soixantaine d'échantillons a ainsi pu être collectée. La liste est disponible en annexe 1.

## IV.2. Les marqueurs

Au départ de cette enquête, il fallait choisir des marqueurs moléculaires adéquats. Tracer les mouvements de plantes en Océanie et tester l'hypothèse tripartite nécessitait notamment de pouvoir discriminer efficacement les groupes génétiques présents en Amérique tropicale. Le pouvoir des marqueurs nucléaires à détecter une structure génétique claire est limité pour une espèce comme la patate douce en raison de la forte hétérozygotie due à son état polyploïde et de son régime de reproduction auto-incompatible. En revanche, les marqueurs chloroplastiques peuvent s'avérer efficaces pour révéler une telle structure génétique et sont des outils moléculaires de choix pour les analyses de phylogéographie (Petit et al. 2005). L'ADN cytoplasmique est généralement à transmission uni-parentale, maternelle chez la plupart des angiospermes (Birky 1988, Provan et al. 2001), et il est non recombinant. Du fait de son caractère haploïde et de son hérédité maternelle, la taille efficace des populations mesurée est théoriquement un quart de celle mesurée avec des marqueurs nucléaires. Ainsi, l'effet de dérive est accentué et la structure génétique a tendance à être plus forte que celle révélée par des marqueurs nucléaires. Aussi le taux de mutation des marqueurs

**Tableau 2 - 2 : Logiciels d'analyses de génétique des populations permettant de traiter des données microsatellites polyploïdes**

Nom du logiciel	Type de données traitées	Analyses statistiques disponibles	Hypothèses de calcul	Référence
<i>GENODIVE</i>	Données autopolyploïdes	Analyse de variance moléculaire, estimation de coefficient de différenciation entre populations, utilisé correspond au nombre d'allèles présents regroupement de type <i>k-means</i> , calcul d'indices de diversité clonal, calcul d'index d'hybridation et tests de Mantel	Pour le calcul des fréquences, le niveau de ploïdie peut être présent en plusieurs copies.	Meirmans et Van Tienderen 2004
<i>SPAGEDI</i>	Données autopolyploïdes	Etude de la structure génétique spatiale des individus cartographiés et / ou des populations cartographiées: l'hypothèse (fausse) que pour un hétérozygote calcul de nombreux indices d'apparentement ou de partiel, tous les allèles ont la même probabilité de différenciation entre individus/populations et d'être présent en plusieurs copies (autocorrélation spatiale, régression linéaire...)	Le calcul des fréquences alléliques est basé sur l'hypothèse (fausse) que pour un hétérozygote partiel, tous les allèles ont la même probabilité d'être présent en plusieurs copies	Hardy et Vekemans 2002
<i>POPDIST</i>	Données autopolyploïdes	Calcul de différentes mesures de distance entre populations quelque soit leur niveau de ploïdie	Le calcul des fréquences alléliques est basée sur l'hypothèse (fausse) que pour un hétérozygote partiel, tous les allèles ont la même probabilité d'être présent en plusieurs copies	Guldbrandtsen et al. 2000; Tomiuk et al. 2009
<i>STRUCTURE</i>	Données autopolyploïdes	Utilise les données de génotypes multilocus pour détecter la structure des populations avec une méthode MCMC. Méthode d'aggrégation de type bayésien. Possibilité de faire de l'assignation à des populations référencées. Détection d'hybrides.	Equilibre de Hardy Weinberg et ségrégation polysomique totale, estimation des fréquences alléliques avec une méthode de type MCMC	Falush et al. 2007
<i>TETRASAT</i>	Données allotétraploïdes	Calcul d'indices de diversité et de différenciation	Enumération de toutes les combinaisons possibles de génotypes (itération) et calcul d'une statistique pour chaque énumération	Markwith et al. 2006
<i>TETRA</i>	Données allotétraploïdes	Calcul d'indices de diversité et de différenciation	Calcule le nombre de copie pour un allèle donné pour un hétérozygote partiel comme le niveau de ploïdie divisé par le nombre d'allèles uniques	Liao et al. 2008
<i>ATETRA</i>	Données allotétraploïdes	Calcul d'indices de diversité et de différenciation	Examine un échantillon aléatoire de toutes les possibilités grâce à une méthode de type MCMC	Van Puyvelde et al. 2010
<i>POLYSAT</i>	Données autopolyploïdes et allotétraploïdes	Interface R pour la gestion de données polyploïdes, et export de fichiers pour de nombreux logiciels, conversion des données en codage binaire, estimation des fréquences alléliques selon deux méthodes, analyse en coordonnées principales...	Possibilité d'estimer les fréquences alléliques en utilisant la méthode développée par Da Silva et al. (2005): estimation itérative (algorithme espérance maximisation) dans un cas de ségrégation polysomique (cas particulier de la ségrégation indépendante des chromosomes)	Clark et Jasieniuk 2011

chloroplastiques est plus faible que celui de l'ADN nucléaire (Gaut et al. 1996). Ces marqueurs permettent donc d'accéder à des patrons génétiques liés à des phénomènes « plus anciens » que ceux accessibles par les marqueurs nucléaires et dans notre cas peuvent être intéressants pour comprendre l'origine des deux groupes génétiques américains (domestication indépendante ou différenciation post-domestication). Enfin, la combinaison de marqueurs nucléaires et chloroplastiques s'avère particulièrement utile pour la reconstitution de l'origine d'un taxon polyploïde et l'identification des différentes lignées parentales en présence.

Deux types de marqueurs chloroplastiques ont alors été utilisés : En premier lieu, la séquence intergénique *rpl32-trnL* (IGS) a été utilisée pour l'étude des relations phylogénétiques entre espèces (chapitre 2). Cette séquence particulière est suffisamment conservée pour une reconstruction d'ordre phylogénétique mais présente aussi un niveau de polymorphisme suffisant pour pouvoir informer sur les relations phylogénétiques à une faible échelle taxonomique (Shaw et al. 2007). En second lieu, des microsatellites chloroplastiques ont été développés pour l'étude de la distribution géographique de la diversité à l'intérieur du compartiment cultivé. Pour les marqueurs nucléaires, nous nous sommes concentrés sur l'analyse de microsatellites nucléaires. De nombreux travaux sur les taxons polyploïdes ont recours à des marqueurs AFLP car les méthodes d'analyse des microsatellites pour une espèce polyploïde sont limitées d'un point de vue théorique et pratique (voir paragraphe suivant). Aussi les techniques d'empreinte génétique permettent d'obtenir un grand nombre de marqueurs et de couvrir largement le génome pour un investissement humain et financier de départ limité. Toutefois, l'amplification des marqueurs AFLP est non spécifique et donc son utilisation est limitée dans le cas de l'amplification d'ADN contaminé - par exemple si le matériel végétal est contaminé par des champignons (ce qui peut arriver assez souvent lorsqu'on travaille avec du matériel séché en conditions tropicales ou équatoriales) - ou encore pour l'amplification de l'ADN dit « sensible » et dégradé des échantillons d'herbiers. Nous avons donc sélectionné une quinzaine de marqueurs microsatellites préalablement développés au CIP.

### **IV.3. Analyses génétiques des polyploïdes à partir de marqueurs microsatellites nucléaires**

La plupart des analyses statistiques qui permettent d'étudier la distribution du polymorphisme neutre de marqueurs microsatellites et de reconstruire l'histoire démographique des populations d'une espèce donnée nécessite au préalable de calculer puis d'estimer les fréquences alléliques à chaque locus (analyse de structuration, de regroupement d'individus ou d'assignation, par exemple). Chez les polyploïdes, l'ambiguïté du nombre de copies présente à chaque locus est un problème majeur (Obbard et al. 2006 ; Clarke et Jasieniuk 2010). Par exemple, si l'amplification de microsatellites chez une espèce tétraploïde pour un locus donné présente les trois allèles suivants ABC, il est impossible de savoir si le génotype exact de l'individu est AABC, ABBC ou encore ABCC et donc de calculer les fréquences alléliques. Esselink et al. (2004) ont proposé une méthode permettant de déterminer le nombre de copies de chaque allèle et basée sur la mesure de l'intensité de fluorescence des différents pics obtenus. Toutefois lorsque le niveau de ploïdie devient important (supérieur à 5), cette méthode s'avère non reproductible et impraticable (Helsen et al. 2009). Il est aussi possible d'estimer les fréquences alléliques d'un polyploïde de façon itérative (en utilisant par exemple, un algorithme de type espérance-maximisation (De Silva et al. 2004) ou des approches bayésiennes (Falush et al. 2007)), mais ces estimations nécessitent de faire des hypothèses sur le fonctionnement des populations et notamment sur le mode de ségrégation du polyploïde. De la même manière qu'il existe un continuum entre la duplication de génomes strictement identiques (qui devraient avoir une ségrégation complètement polysomique) et celle de génomes très différenciés (qui devraient avoir une ségrégation de type disomique), de nombreux polyploïdes ont des patrons de ségrégation intermédiaires (Obbard et al. 2006) qu'il peut être difficile de « modéliser ».

Au vu de ces difficultés théoriques et pratiques, peu de logiciels d'analyse de génétique des populations sont finalement capables de prendre en compte des données microsatellites polyploïdes (Clarke et Jasieniuk 2010), ou sont restreints à un type de ségrégation, limitant dès lors le panel d'analyses envisageables (cf. Tableau 2-2). Ces limitations analytiques ont souvent conduit les

chercheurs à coder leurs données microsatellites polyploïdes comme des marqueurs dominants, c'est-à-dire à utiliser un codage binaire en 0/1 (0 pour l'absence de l'allèle et 1 pour sa présence). Ce type de codage représente une perte d'information importante (perte de l'information sur l'hétérozygotie) mais permet d'accéder à de nombreux autres logiciels et méthodes statistiques capables de traiter de telles données (voir Bonnin et al. 2007 pour une revue récente). Récemment, la publication d'un programme R « Polysat » a facilité grandement la gestion et la conversion des données microsatellites polyploïdes (Clarke et Jasieniuk 2010). Néanmoins, il reste difficile d'évaluer les conséquences d'un tel codage sur la précision et l'exactitude des méthodes utilisées ensuite.

Aussi, d'autres méthodes d'analyses de la structuration de la diversité sont basées sur le calcul d'indices de similarité (ou de dissimilarité) entre individus puis sur la représentation de cette diversité au travers de méthodes factorielles (qui recherchent une représentation globale de la diversité et chassent les effets individuels) ou de représentations arborées (qui au contraire cherchent à représenter les relations individuelles). Il existe de nombreux indices de similarité possibles (indices de Sokal et Michener, de Dice, de Jaccard...). Certains indices prennent en compte à la fois les « présences et absences » partagées (indice de Sokal et Michener), d'autres ne considèrent que les « présences partagées » (indice de Jaccard ou indice de Dice, par exemple) dans le calcul de la similarité. Selon le type de marqueurs utilisés (dominant, codominant), le niveau de ploïdie, mais aussi le niveau d'hétérozygotie des individus en question, les différents indices apparaissent plus ou moins pertinents (Perrier et al. 1999 ; Kosman et Leonard 2005). Le choix de l'indice pour les taxons polyploïdes s'avère particulièrement délicat et aucune étude systématique ne permet d'éclairer ce choix (Kosman et Leonard 2005). Les indices considérant uniquement les présences de bandes sont généralement utilisés (comme par exemple l'indice de Lynch ou encore de Jaccard). Toutefois, il est en général préférable d'utiliser plusieurs types d'indices et de comparer les patrons de diversité observés.

Compte tenu des difficultés énoncées précédemment, l'analyse des données polyploïdes est délicate et la précision des méthodes difficiles à évaluer. Il est donc essentiel de croiser les approches. Les méthodes d'analyses utilisées dans cette thèse sont explicitées dans les chapitres suivants.

Problématique de recherche	Echantillonnage	Collections	Marqueurs moléculaires	Analyses	Echelle géographique	Chapitre
<b>Caractériser la structure de la diversité génétique de la patate douce sur son aire d'origine</b> <i>La structure de la diversité des cultivars dépend-elle d'événements multi-locaux de domestication?</i>	329 cultivars	CIP	7 SSRs chloroplastiques, 13 SSRs nucléaires	diversité et différenciation génétique, réseau d'haplotypes, regroupement K-means	Amérique tropicale (du Mexique au Pérou)	chap. 2
<b>Inférer l'origine botanique et géographique de <i>Ipomoea batatas</i></b> <i>Identifier les parents sauvages impliqués dans la formation du génome hexaploïde de la patate douce, localiser le ou les foyers de domestication, établir la relation temporelle entre polyploïdisation et domestication</i>	297 accessions; 6 espèces de la série Batatas ( <i>I. batatas</i> , <i>I. trifida</i> , <i>I. triloba</i> , <i>I. leucantha</i> , <i>I. tiliacea</i> , <i>I. tabasacana</i> )	CIP, NIAS et USDA	1 séquence chloroplastique (rpl32-trnL IGS); 1 séquence nucléaire (ITS); 8 SSRs nucléaires	arbres phylogénétiques (bayésien et de distance), regroupement bayésien, regroupement K-means et c-means	Amérique tropicale (du Mexique au Pérou)	chap. 3
<b>Comprendre les processus à l'origine de la création d'un centre de diversité secondaire: le cas de la patate douce en Nouvelle-Guinée</b> <i>Interactions des facteurs historiques, biologiques et culturelles</i>	417 cultivars	NARI, CIP	6 SSRs chloroplastiques, 11 SSRs nucléaires	réseaux d'haplotypes, diversité et différenciation génétique, assignations (analyse discriminante et regroupement bayésien)	Ile de Nouvelle Guinée et archipel Bismarck, Amérique tropicale	chap. 4
<b>Retracer l'histoire d'introduction de la patate douce en Océanie</b> <i>Apporter des preuves génétiques à l'hypothèse "tripartite" développée par les ethnobotanistes, archéologues et linguistes pour expliquer la distribution de la patate douce dans le Pacifique.</i> <i>Comparaison régionale des patrons de diversité et des processus de diversification sous-jacents</i>	1245 cultivars: 1188 contemporains 57 spécimens d'herbiers du 17 <sup>ième</sup> au 20 <sup>ième</sup> siècle	<i>ex situ</i> CIP, NARI, NICS, SDR, CTRAV herbiers BISH, BM, K, L	6 SSRs chloroplastiques, 11 SSRs nucléaires	réseaux d'haplotypes, diversité et différenciation génétique, assignations (analyse discriminante et regroupement bayésien)	Océanie et ses "marges" continentales (Asie du Sud-Est et Amérique tropicale)	chap. 5

# *CHAPITRE 3*

*Diversité cultivée  
2) Amérique tropicale*

*Article : Caractérisation de la structure de la  
diversité génétique de la patate douce dans son aire  
d'origine*

*Combining chloroplast and nuclear microsatellites  
to investigate origin and dispersal of New World  
sweet potato landraces*

**1. Title:** Combining chloroplast and nuclear microsatellites to investigate origin and dispersal of New World sweet potato landraces

**2. Authors' names:** C. Roullier<sup>1,2</sup>, G. Rossel<sup>3</sup>, D. Tay<sup>3</sup>, D. McKey<sup>2</sup> and V. Lebot<sup>1</sup>

**3. Full postal address:**

<sup>1</sup> CIRAD Avenue Agropolis, 34398 Montpellier Cédex 5, France

<sup>2</sup> Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175 CNRS, 1919 route de Mende, 34293 Montpellier Cédex 5, France

<sup>3</sup> International Potato Center (CIP), Lima, Peru

**4. Keywords:** Chloroplast microsatellites, Domestication, Genetic diversity, *Ipomoea batatas*, Phylogeography.

**5. Corresponding author:**

Roullier Caroline, CIRAD Avenue Agropolis, 34398 Montpellier Cédex 5

France. Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175, 1919 route de Mende, 34293 Montpellier Cédex 5, France; E-mail address: caroline.roullier@cefe.cnrs.fr.

**6. Running title:** New World sweet potato genetic diversity

**7. Article published in Molecular Ecology (2011).**

## ABSTRACT

We analyzed a representative collection of New World sweet potato landraces (329 accessions from Mexico to Peru) with both chloroplast and nuclear microsatellite markers. Both kinds of markers supported the existence of two geographically restricted genepools, corresponding to accessions from the north-western part of South America and accessions from the Caribbean and Central America region. Our conservative cpSSRs markers revealed that the divergence between the two haplotype groups is associated with numerous mutation events concerning various markers, supporting the idea that this divergence may be ancient, predating domestication. For both kinds of markers, we found no significant difference in diversity between the two genepools and detected region-specific alleles in both groups. Previous studies have favoured the hypothesis of a single domestication of this crop. Our analysis suggests at least two independent domestications, in Central / Caribbean America and in the north-western part of South America. Sweet potato was then dispersed from these centres throughout tropical America. Comparison of nuclear and chloroplast data suggests that exchanges of clones and sexual reproduction were both important processes in landrace diversification in this clonally propagated crop. Our analysis provides useful tools for rationalizing the conservation and use of sweet potato germplasm collections.



## INTRODUCTION

Molecular studies have greatly enhanced our understanding of where, when and how crop plants arose (Matsuoka *et al.* 2002; Sanjur *et al.* 2002; Spooner *et al.* 2005; Chácon *et al.* 2005, Londo *et al.* 2006). Despite its importance as a staple food, ranked seventh among world crops with an annual production of around 9 Mt and a cultivated area of 110 M ha, (Lebot, 2010), sweet potato (*Ipomoea batatas* (L.) Lam. Convolvulaceae) has been relatively neglected and many points of its evolutionary history from domestication to dispersal remain unsolved.

As currently circumscribed, *I. batatas* is a hexaploid species, vegetatively propagated by farmers but having a sporophytic incompatibility reproduction system, and including mainly domesticated types. Following a taxonomic revision of the *I. batatas* complex, Austin (1987) proposed that natural hybridisation between *I. trifida* and *I. triloba* generated the wild ancestors of *I. batatas* somewhere between the Yucatan peninsula in Mexico and the Orinoco basin in Venezuela. Kobayashi (1984), however, suggested that *I. trifida* is an autopolyploid complex with different ploidy levels from 2X to 6X, from which *I. batatas*, the domesticated form, has derived. It is therefore possible that *I. batatas* could have been generated by natural hybridisations involving several wild species, rather than descending from a single wild ancestor (Lebot 2009). Analysis of molecular markers, including RAPD (Jarret and Austin 1994), microsatellites (Buteler *et al.* 1999), ISSR and cpRFLP (Huang and Sun 2000) and AFLP (Huang *et al.* 2002), a phylogenetic analysis based on beta-amylase gene sequences (Rajapakse *et al.* 2004) and more recently, a cytogenetic study (Srisuwan *et al.* 2006), have all confirmed the close relationship between *I. batatas* and *I. trifida*. Because the natural distribution of *I. trifida* extends from northern Peru to Mexico (Shiotani *et al.* 1990), it is logical to assume that the domestication of sweet potato occurred within this vast geographical region, but how, where and how many times remain open questions. A genetic study based on AFLPs indicated that the highest genetic diversity was found in Central America, whereas the lowest was in the Peru-Ecuador region (Zhang *et al.* 2000). No region-specific profiles were identified but a significant regional differentiation between Peruvian and Ecuadorian cultivars and other American cultivars was demonstrated. For these authors, the results confirm previous molecular studies and support the hypothesis of Central America or Meso-America as the primary centre of diversity and the most likely centre of origin of domesticated *I. batatas*. Sweet potato was then dispersed to the Peru-Ecuador region, where a secondary centre of diversity formed. Another genetic study using RAPD markers and 74 sweet potato accessions from its worldwide range confirmed the existence of two distinct gene pools in American sweet potato landraces (Gichuki *et al.* 2003).

Numerous important crops were domesticated in the Americas, including maize (*Zea mays*), potato (*Solanum tuberosum*), common bean (*Phaseolus vulgaris*) and cassava (*Manihot esculenta*) (Pickersgill 2007). Some plants were domesticated once from their wild progenitor in a restricted geographical area, such as maize in Mexico (Matsuoka *et al.* 2002, Ranere *et al.* 2009, Van Heerwaarden *et al.* 2011). Quite frequently, however, species of the same genus were domesticated independently in different areas, as in *Cucurbita* (Sanjur *et al.* 2002). More rarely, the same wild ancestral species was domesticated independently in different parts of its distributional range, as in *Phaseolus vulgaris* (Chacón *et al.* 2005). Plants domesticated once and that have dispersed quickly and widely from their centre of origin, as peach palm (*Bactris gasipaes*, Arecaceae) landraces (Clement *et al.* 2010) or maize (Vigouroux *et al.* 2008) exhibit clear genetic structure among populations, resulting of isolation by distance, genetic drift and local adaptation. Structure of diversity among landraces results from both domestication and diffusion processes. Appropriate markers and sampling designs are then required to distinguish multiple independent domestication events from scenarios involving diffusion from a single domestication event.

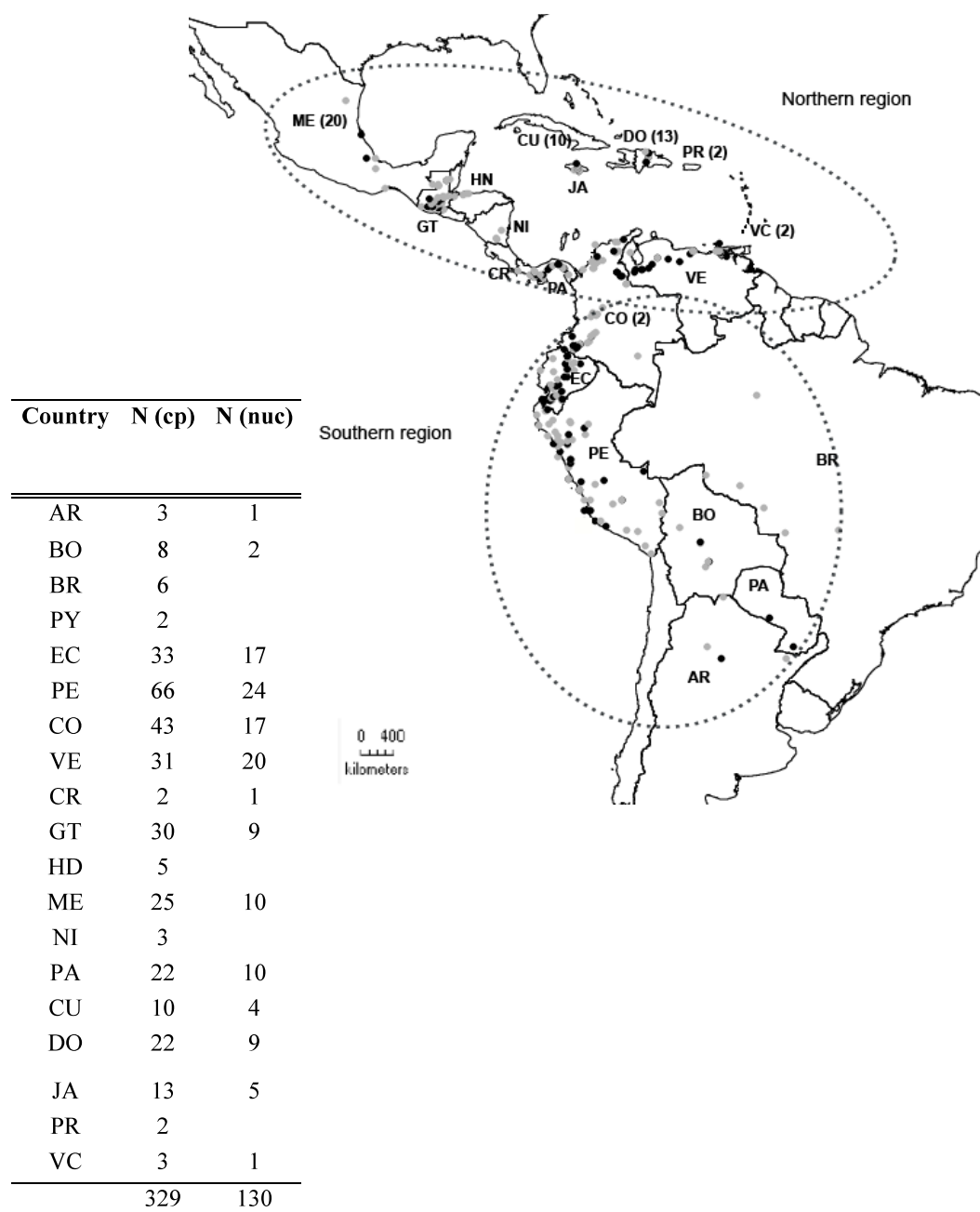
Chloroplast microsatellites (cpSSRs) are useful tools for evolutionary studies at low taxonomic levels, as in crop domestication studies, because they often exhibit geographically informative polymorphism among populations and also among closely related species (Provan *et al.* 2001). They have, for example, permitted the identification of progenitors and of the geographical origin of domestication (Wills *et al.* 2006; Xu *et al.* 2002; Imazio *et al.* 2006). A set of conserved primer pairs that flank cpSSRs have been identified from a number of angiosperm species (Bryan *et al.* 1999; Weising and Gardner 1999), simplifying the use of chloroplast markers for plants for which little sequence information is available. Combining nuclear SSRs, which are more polymorphic and

bi-parentally inherited, with the haploid cpSSRS, non-recombinant and usually maternally inherited, permits access to different parts of the history of populations: nuclear SSRs are suitable for studying recent and local evolutionary processes while cpSSRs are effective indicators of population subdivision and differentiation for tracing ancient divergences.

All attempts so far to structure the existing morphological variation found within cultivated *I. batatas* into meaningful varietal groups have failed because variation appears to be continuous. This is a serious constraint for germplasm curators and especially for breeders, who therefore must rely on comprehensive studies characterizing gene pools to identify groups potentially differing in functionally important genetic diversity (Lebot, 2010). All genetic studies of American sweet potato landraces have suffered from very restricted sampling and from the use of markers that are to identify consistent gene pools and determine their origin inadequate (Zhang *et al.* 2000, 2001, Gichuki *et al.* 2003). To investigate genetic diversity in American sweet potato and to analyze the contributions of domestication and post-domestication dispersal in explaining the patterns observed, we examined a comprehensive sample of landraces (329 sweet potato landraces from Mexico to Peru) using both, chloroplast and nuclear microsatellites. We present new insights into the evolutionary history of the crop.

**Figure 3-1: Geographic distribution of CIP sweet potato accessions used in the present study.**

No geo-referenced accessions are indicated per country between brackets. For the well-referenced ones, circles (grey and black) indicate accessions genotyped for chloroplast SSRs. Grey circles indicate the subsample genotyped also for nuclear SSRs. The table provides the number of accessions per country and geographical area for samples genotyped with cpSSRs (Ncp) and nuclear SSRs (Nnuc).



## MATERIALS AND METHODS

### *Analysis with chloroplast markers*

#### Sampling area and plant materials

The sampling area extends from Mexico to Peru with collection sites ranging in elevation from 0 to 2850 meters. Among the 4615 sweet potato landraces maintained in the international collection at the International Potato Center (CIP, Lima, Peru) 329 accessions with complete passport data were selected among to cover the entire pre-Columbian range of sweet potato cultivation (Figure 3-1).

Healthy young leaves collected from accessions maintained in screen houses or in vitro were freeze dried at CIP and stored at room temperature until use. All extractions were conducted using the Qiagen 96 Plant kit for lyophilised tissues (Hilden, Germany).

#### Amplification of chloroplast microsatellites

Overall, 44 universal primers (Bryan *et al.* 1999; Weising and Gardner 1999) were tested on a worldwide subsample of 12 sweet potato landraces from the CIP genebank and only four primers (ccmp2, NTCP 18, NTCP 28, NTCP 26) showed polymorphism. We then developed additional specific primers for *Ipomoea* species, using the complete chloroplast genome sequence of *I. purpurea*, available from the National Center for Biotechnology Information (NCBI) genebank, to identify all (A)<sub>n</sub> and (T)<sub>n</sub> repeats with  $n \leq 10$ . Forty microsatellites were characterized and we then tested amplification and readable polymorphism on the same worldwide subsample. Three additional markers were selected (Ibcp 5, Ibcp 8 and Ibcp10). Thus a total of seven chloroplast SSRs were used in our study (Table 3-1).

All loci were amplified independently using Multiplex PCR Taq (Qiagen) in a final volume of 10 $\mu$ l. The following program was conducted, using a PTC-100 Thermocycler (MJ Research, Waltham, MA, USA) 15 min. at 95°C / 30 cycles of 30 sec at 94°C, 1 min. 30 sec at T<sub>m</sub> primer (Table 3-1), 1 min. at 72°C / 30 min. at 60°C. Ninety-six-well plates were used for amplification and three wells per plate were devoted to negative (one well) and migration controls (two wells). Because cpSSRs have just a one bp mutation step, we checked repeatability by genotyping twice 50 randomly selected samples in the different plates. Genotyping was performed on an ABI 3130 sequencer (Applied Biosystems, Foster City, CA, USA) and all genotypes were visually checked by two independent investigators using GENEMAPPER 3.0 software (Applied Biosystems). To verify whether the observed variations were due to changes in mononucleotide repeats or to other mutation events, all size variants detected in this study were sequenced.

#### Analysis of haplotypes

Multilocus haplotypes were defined using GENALEX 6 (Peakall and Smouse 2006). To infer haplotypes relationships, we performed a factorial analysis based on a distance matrix between haplotypes using the Darwin software (version 5.0, Perrier and Jacquemoud-Collet 2006). The established distance corresponds to the simple matching dissimilarity index for haploid data. A median joining network analysis was also computed with Network 4.5.6.1 (<http://www.fluxusengineering.com.sharenet.htm>). One-bp mutations were given a weight of one.

Both analysis suggested the existence of two divergent genepools associated to two geographical regions, the North western part of South America (with accessions from Peru, Ecuador and South Columbia), and the Central / Caribbean part of America (with accessions from Mexico to North Columbia), thereafter referred to as Southern region and Northern region, respectively (Figure 3-1). We realized an AMOVA with Arlequin software 3.1 (Excoffier *et al.* 2005) to test statistically this clear-cut genetic structure. We recognized two groups, corresponding to accessions from the Southern and Northern regions, within each we classed geographical “populations” corresponding to groups of landraces from a given country.

The total number of cpSSR haplotypes and the number of private haplotypes were then calculated for each country and summarized over the Southern and the Northern regions, using GENALEX 6 (Peakall and Smouse 2006). To account for differences in sample size, rarefaction analysis of haplotype diversity was performed using CONTRIB (Petit *et al.* 1998), for countries from which more than five accessions were available.

### *Analysis with nuclear SSRs*

#### Plant materials

**Table 3-1: Characterization of the seven cpSSR markers used in our study and the polymorphism observed for them.**

Locus Name	Primer sequences (5'-3')	Allele size range	T (°C)	Number of alleles	Gene diversity
ccmp 2	F: GATCCCGGACGTAATCCTG R: ATCGTACCGAGGGTTCGAAT	199-202	60	4	0.625
NTCP 18	F: CTGTTCTTTCCATGACCCCTC R: CCACCTAGCCAAGCCAGA	194-198	60	4	0.104
NTCP 28	F: TCCAATGGCTTTGGCTA R: AGAAACGAAGGAACCCAC	178-183	58	5	0.509
NTCP 26	F: GCAATTGCAATGGCTTCTTTA R: TTTATGTTTCGGTGGAAATCACA	175-178	60	4	0.502
Ibcp 5	F: GCTCTCACGCTCAATTACTTA R: ATGCTTAATTGACGACCTGT	242-246	55	4	0.486
Ibcp 8	F: AATAAGTACTTGGCCGTGAA R: CGATTCAAGTAGGCAAAGAG	195-197	55	3	0.313
Ibcp 10	F: ATATAAGGGGCCATTTTAGG R: ACGATAGAGGAGAAGGTTCC	199-201	55	3	0.490
<i>Mean</i>				<i>3.857</i>	<i>0.4327</i>

130 accessions were selected from the previous sampling exercise to provide representative coverage of the sampling area and the two described chloroplast genepools (Figure 3-1).

#### Amplification and scoring

Thirteen nuclear microsatellites designed at the CIP (Solis *et al.* unpublished data) were selected for the present study based on their reliable amplification profiles in Sequencer and their polymorphism (Table 3-2). The amplification protocol used was the same as that used for cpSSRs. Because nuclear SSR profiles are often difficult to read in this hexaploid species, all samples were genotyped at least twice. All profiles that were not reproducible were counted as missing data. Genotyping was performed on an ABI 3130 sequencer, and all profiles were eye-checked by two independent investigators using GENEMAPPER 3.0 software.

For all analyses that required calculations of distances, our microsatellite data were coded in a binary manner similar to that used for AFLP data (0 = absence, 1 = presence of a given allele). Over all loci, 130 alleles were obtained; we thus used 130 binary variables.

#### Patterns of diversity

As sweet potato varieties are clonally propagated, we estimated the number of unique multilocus genotypes (MLGs), and then identified sets of clones. To identify clonal lineages, we followed the procedure described by Arnaud-Haond *et al.* (2007). We screened each MLG pair presenting extremely low genetic distances. These pairs accounted for a small peak (at very low values) in the frequency distribution of distances, making this distribution bimodal rather than unimodal. Darwin software was used to obtain the frequency distribution of pairwise distances between our 99 MLGs. Because our data set on this hexaploid species does not permit knowing exact allelic frequencies, the second step, estimating the likelihood that slightly distinct MLGs were derived from distinct recombination events, was impossible. To compare the number of alleles among groups of different sizes, we used a rarefaction method (Petit *et al.* 1998, Vigouroux *et al.* 2008). The 'R' statistical software was used to calculate the average number of alleles for each locus in each group as the mean of 1000 re-samplings at the smallest group sample size (i.e. the number of plants assigned to the smallest group). Numbers of alleles per locus and per group were compared using Wilcoxon signed-rank tests with 'R'.

#### Genetic structure as estimated using nuclear SSRs

First, we used a non-model-based algorithm which has already been successfully applied in phylogeographic studies based on AFLP (Burnier *et al.* 2009; Arrigo *et al.* 2010). Because this analysis does not allow missing data, our subsample was reduced to 101 accessions. Non-hierarchical K-means clustering (Hartigan and Wong 1979) was performed using R, running scripts developed by Arrigo *et al.* (2010) and kindly provided by the authors. This technique assigns individuals to a defined number of genetic groups (hereafter K) in order to maximize the intergroup variance, measured here as the inertia. We performed 100 000 independent runs (i.e. starting from random points) for each assumed value of K (ranging between 1 and 10) and recorded the intergroup inertia of each run. We then explored the geographical composition of the genetic clusters from K = 2 to K = 10, corresponding to the replicate that maximizes the inter-group inertia.

A factorial analysis was performed with the software Darwin 5. A genetic distance matrix was obtained by averaging over all loci Sokal and Michener distances computed for each locus. Mantel tests were also performed to compare the matrix of genetic distances (Sokal and Michener) between individual plants with a matrix of geographical distances calculated from latitude and longitude of each accession by DIVA-GIS 7.3 (Hijman *et al.* 2002). The correlation was tested over the whole sampling area, as well as within the Southern and Northern region.



**Table 3-2: Characterization of the nuclear SSRs used in our study.**

Locus Name	T (°C)	Primer sequences (5'-3')	SSR Motif	Range Size
J263	57	F: CTCTGCTTCTCCTGCTGCTT R: GTGCGGCACTTGTCTTTGATA	(AAC)6	156-171
J315E	57	F: TAGGTGTGTTTATGGGAGATT R: GGGACTTGACTTTCATTATTAC	(TCAG)5	232-243
J522A	57	F: ACCCGCATAGACACTCACCT R: TGACCGAAGTGTATCTAGTGG	(CAC)7	208-246
J116a	57	F: TCTTTTGCATCAAAGAAATCCA R: CCTCAGCTTCTGGGAAACAG	(CCT)7	187-227
Ib297	58	F: GCAATTTACACACAAACACG R: CCCTTCTTCCACCACTTTCA	(CT)13	129-167
J10A	58	F: TCAACCACTTTCATTCACTCC R: GTAATTCCACCTTGCGAAGC	(AAG)6	173-209
J206A	58	F: ATCAGGGAGAGAGGACAGTAA R: TAGGCAAACCATAAACAGAGA	(GAT)6	103-121
J130	60	F: CCAGTTGTGGCGGAGAAAG R: GGATTCTTGTGAGGGACAACA	(GAA)8	150-159
J1809E	60	F: CTTCTCTTGCTCGCCTGTTC R: GATAGTCGGAGGCATCTCCA	(CCT)6	120-135
IbR16	60	F: GACTTCCTTGGTGTAGTTGC R: AGGGTTAAGCGGGAGACT	(GATA)4	194-213
IbC5	60	F: CCACAAAAATCCCAGTCAACA R: AGTGGTCGTCGACGTAGGTT	(AAG)8	89-109
J544b	60	F: AGCAGTTGAGGAAAGCAAGG R: CAGGATTTACAGCCCCAGAA	(TCT)6	174-194
IbS11	60	F: CCCTGCGAAATCGAAATCT R: GGACTTCCTCTGCCTTGTTG	(TTC)10	218 -248

## RESULTS

### *Analysis of haplotypes*

All loci harboured polymorphism, with a mean of 3.86 ( $\pm$  0.69) alleles per locus, resulting in an average haploid genetic diversity of 0.432 ( $\pm$  0.171) (Table 3-1). The sequences of the cpSSRs revealed that differences in length among the PCR products were due to differences in the number of mononucleotide repeats. No other length mutations, such as insertions or deletions, were detected in the sequenced accessions. In total, we identified 24 haplotypes (Table 3-3), of which the five most common (H2, H6, H13, H14 and H16) occurred at frequencies of 0.09 to 0.23 and the 19 others at frequencies less than 0.03, while eight haplotypes were found only once.

Genetic relationships between haplotypes were first investigated by a factorial analysis. Sweet potato haplotypes formed two main clusters defined by the first axis (Figure 3-2a), which explained 72.0 % of the total variance; the second axis explained 7.8 %. Each group was compact, with some outlying haplotypes (10-7, 8-4-5). We then used a network approach to improve our understanding of genetic relationships between different haplotypes. The minimum spanning tree distinguished two groups (Figure 3-2b) similar to those revealed by the factorial analysis. These groups (groups 1 and 2) were separated from each other by several mutation events. Each group exhibited a star-like topology with rare and geographically restricted variants radiating from the most widely distributed and commonly occurring haplotypes. H10 and H21 were the only two haplotypes that clearly diverged from their respective parent groups by several mutation events. This tree organisation supports the hypothesis of an ancient divergence between the two main groups, within each of which recent diversification occurred by mutation events.

The distribution of haplotypes showed strong geographical structure. The Southern and Northern regions were differentiated on the basis of their haplotype compositions: group 1 haplotypes were mainly encountered in the Northern region, whereas group 2 haplotypes were predominant in the Southern region (Table 3-4). Notable exceptions to this clear-cut pattern concern specifically the most common haplotypes (Figure 3-2b and Table 3-4), suggesting either long-distance dispersal events or misappropriation of samples to geographic region in passport data. The AMOVA analysis confirmed this genetic structure: a large part of the total molecular variance (35.27%,  $P < 0.0001$ ) is attributed to the among regions variation, whereas the among populations within region variation accounts for 9.60% ( $P < 0.001$ ). The maximum of diversity comes from within-region variation (55.3%,  $P < 0.001$ ) illustrating the fact that what we defined as “populations” are heterogeneous, containing in particular haplotypes from both genetic groups.

The number of haplotypes per country ranged from 1 to 10 (mean 4). Sixteen haplotypes are found in the Northern region of which 7 are private. Fifteen are present in the Southern region of which 7 are private. The highest haplotype diversity (0.83) was found for Ecuadorian landraces and the lowest (0.61) for those from the Dominican Republic (Table 3-4). Peru and Columbia possess most of the private haplotypes.

### *Population structure revealed with nuclear SSRs*

The hierarchical K-means clustering was run for 101 accessions (a subsample of unique multilocus genotypes with no missing data). For  $K = 2$ , the first group included 40 accessions, mainly accessions from the Southern region, but also including five from the Northern region. The second group was almost exclusively represented by 62 samples from the Northern region, including only six accessions from the Southern region (two from Ecuador, two from Peru and two from southern Columbia) (Figure 3a and 3b). The composition of clusters for higher  $K$  values allows us to explore more subtle substructures. We found that, in general, accessions from the Southern region clearly clustered in two regional groups, the southern-Columbia / Ecuador group and the Peruvian group. In contrast, accessions from the Northern region were divided into different subgroups with accessions from distinct geographical areas.

Results of factorial analysis based on Sokal and Michener distance for the 130 subsamples were largely congruent with the non-hierarchical clustering for  $K = 2$ . The two first axes explained only 15.8% of the total variance. The first axis distinguished accessions of the Southern region from those

**Table 3-3: Multilocus haplotypes defined with seven chloroplast markers (ccmp2, ntcp 18, ntcp 28, ntcp 26, ibcp 10, ibcp5 and ibcp8)**

Haplotype number	ccmp2	ntcp18	ntcp28	ntcp26	ibcp10	ibcp5	ibcp8
1	199	194	179	176	199	244	195
2	199	195	179	176	199	244	195
3	199	195	180	176	199	244	195
4	200	195	178	177	201	243	197
5	200	195	178	177	201	244	197
6	200	195	179	176	199	244	195
7	200	195	179	176	199	246	195
8	200	195	179	177	201	243	197
9	200	196	179	176	199	244	195
10	200	198	183	176	199	243	196
11	201	195	178	177	200	242	196
12	201	195	178	177	200	243	196
13	201	195	178	177	201	243	197
14	201	195	178	177	201	244	197
15	201	195	178	178	201	243	197
16	201	195	179	176	199	244	195
17	201	195	179	176	199	246	195
18	201	195	179	177	199	244	195
19	201	195	179	177	201	243	197
20	201	195	180	177	201	243	197
21	201	195	182	177	201	244	197
22	201	196	179	175	199	244	195
23	201	196	179	176	199	244	195
24	202	195	179	176	199	244	195

of the Northern one. The second axis distinguished Peruvian accessions from Ecuadorian-southern Colombian ones (Figure 3-4). Geographical sub-structure of accessions from the Northern region was also not clearly characterized by the factorial analysis.

The correlation between genetic and geographic distances was significant for the entire sample ( $r = 0.212$ ,  $P = 0.001$ ), for the Southern region ( $r = 0.203$ ,  $P = 0.001$ ) and for the Northern region ( $r = 0.203$ ,  $P = 0.003$ ).

*Comparison between genetic structures revealed by chloroplast and nuclear SSRs*

Phylogeographic patterns revealed by chloroplast and nuclear data were globally congruent: Accessions from the Southern region were mainly associated with cp haplotype group 1 and nuclear cluster K1 (Figure 3-4). Accessions from the Northern region were mainly of haplotype group 2 and nuclear cluster K2. Both kinds of markers supported the existence of two geographically restricted genepools. However, 24 accessions possessed group 1 haplotypes and nuclear diversity placing them in cluster K2. Most of these (19) originated from the Northern region. Similarly, 13 accessions possessed group 2 haplotypes and nuclear diversity associated with cluster K1. Most of these originated from the Southern region.

*Pattern of nuclear SSR genetic diversity*

Within the individuals genotyped with nuclear SSRs and without missing data, we found 101 unique multilocus genotypes. We detected six sets of two or three replicates of the same MLG per set. All replicates were locally restricted. Following the criteria of Arnaud-Haond *et al.* (2007), we identified seven clonal lineages that probably correspond to groups of variants derived by somatic mutation. Most accessions from a given putative clonal lineage were quite restricted in distribution.

The rarefied allele number of the Northern region genepool was slightly higher than that of the Southern region genepool, but the difference was not significant ( $P = 0.55$ ). Together, the Southern genepool exhibited 13 private alleles while the Northern one showed 17 (Table 3-5).

N represents the total number of accessions, Ht the number of different haplotypes identified, Ht(1) the number of different haplotypes from group 1, Ht(2) the number of different haplotypes from group 2 and hr the haplotype diversity as calculated by rarefaction analysis. Group 1 Haplotypes are underlined in light grey and group 2 haplotypes in dark grey. ND means not determined.

Geographic region		Haplotype																					N	Ht	Ht(1)	Ht(2)	Hp			
		H1	H2	H3	H6	H7	H9	H10	H16	H17	H18	H22	H23	H24	H4	H5	H8	H11	H12	H13	H14	H15						H19	H20	H21
Northern region	Mexico		11	3	5		2		1		1									2						25	7	1	6	1
	Nicaragua		2		1																					3	2		2	
	Panama		6						8			4	1								3					22	5	1	4	1
	Guatemala		14		10					2										2	2					30	5	2	3	1
	Honduras		3		2																					5	2		2	
	Costa Rica		1																		1					2	2	1	1	
	Cuba		1	3	4																2					10	4	1	3	
	Dominican Republic		13	1	4																4					22	4	1	3	
	Jamaica		10	2																	1					13	3	1	2	
	Puerto Rico		2																							2	1		1	
	Saint Vincent	1	1																		1					3	3	1	2	1
	Venezuela		8		11					5					1						6					31	5	2	3	
	North Columbia		1		1	3		1	10				1	4							3					24	8	1	7	3
	Total Northern region		1	73	9	38	3	2	1	24	2	1	4	2	4	1					2	25					137	16	3	13
Southern region	South Columbia				10				1			2		2						4						19	5	2	3	
	Ecuador		1		2				3			2		1		2				9	10	3				33	9	5	4	1
	Peru		1									4		1	2	2	2			21	28		4	1		66	10	8	2	4
	Paraguay								1												1					2	2	1	1	
	Argentina				1																2					3	2	1	1	
	Bolivia				2																	2			1	8	4	3	1	2
	Brazil		3							1											2					6	3	1	2	
Total Southern region			5		15				6			8		4	2	4	2	1	34	47	3	4	1	1		192	15	11	4	3

## DISCUSSION

The maintenance of *ex situ* collections of sweet potato is a laborious undertaking. Major concerns are the elimination of duplicates and the elaboration of a core collection. Our results provide tools for achieving these goals, permitting characterization of sweet potato genetic diversity. Previous studies, notably on potato, have demonstrated the importance for breeders of considering the cytoplasmic genetic base (Provan *et al.* 1999). In our case, cytoplasmic markers clearly revealed the existence of two different gene pools, which were less well discriminated by nuclear markers. It is entirely possible that different sets of functional genetic diversity may be contained in the two groups thus defined, underlining the importance for breeders of taking cytoplasmic markers into account.

Previous studies based on AFLP and RAPD markers had already suggested the existence of two distinct genepools in tropical America (Zhang *et al.* 2000; Gichuki *et al.* 2003), one associated with the Peru-Ecuador region and the other with the Caribbean and Central America. However, whereas these authors focused solely on the hypothesis of a single origin of domestication in the latter region, followed by dispersal to Peru and Ecuador, our results suggest the pertinence of considering the alternative hypothesis of independent domestications in Meso-America and in north-western South America. We here discuss these two scenarios for the origin of domesticated sweet potato.

### *Existence of two distinct genepools and the origin of cultivated sweet potatoes*

Our conservative cpSSRs markers revealed that the divergence between the two haplotype groups is associated with numerous mutation events concerning various markers, supporting the idea that this divergence may be ancient. Because mutation model(s) and mutation rates of chloroplast microsatellites are not well understood, and furthermore are quite difficult to estimate with clonally reproducing organisms, they are not suitable markers for inferring demographic events, including estimation of the timing of divergence between the two groups identified in this study (Provan *et al.* 2001). However, the congruent and strong phylogeographic signal from both cpSSRs and nuclear SSRs supports the idea that this divergence may be associated with ancient geographical differentiation of wild ancestral populations, predating sweet potato domestication.

### Scenario 1: a single ancient origin

In the previous studies (Zhang *et al.* 2000; Gichuki *et al.* 2003), the hypothesis that domestication originated somewhere in the northern Neotropics (Caribbean, Central America or extreme northern South America) was favoured by the lower level of sweet potato genetic diversity found in the Peru-Ecuador gene pool, the absence of region-specific AFLP or RAPD profiles, and the greater diversity of wild *Ipomoea* species in Central America (Austin, 1987). According to this hypothesis, the Peru-Ecuador group would have diverged from the northern genepool during dispersal, accompanied by reduction of diversity, and by adaptation driven by new environments and human selection.

Under such a scenario, domesticated plants exhibiting group 2 chloroplast haplotypes, subsequent to their diffusion into the Southern region, could have captured group 1 haplotypes from divergent local wild populations to obtain the observed pattern of chloroplast genetic diversity. Such putative wild/cultivated hybridization events must have occurred very early during the diffusion of sweet potato into the Southern area, as group 1 haplotypes strongly predominate in that region.

As a result of founder effects, it is likely that at least some loss of genetic diversity occurred during the expansion of sweet potato from a putative single domestication centre in Meso-America. Contradicting this expectation, we found no significant difference in diversity between the Northern and Southern pools and detected region-specific alleles in both groups, appearing with frequencies < 0.2. The opposite scenario, with initial domestication in South America and diffusion into Caribbean-Central America, is therefore equally plausible. In either case, gene flow from multiple hybridization events with local wild populations could have maintained and/ or enhanced allelic diversity. The existence of two haplotypes, H10 and H21, very distant from the common ones, indicates that wild/cultivated gene flow probably occurred. Such interspecific hybridization has already been recorded in natural conditions (Austin, 1977; Austin, 1987) and has been used in sweet potato genetic improvement (Orjeda *et al.* 1991; Shiotani *et al.* 1990). However, incompatibility among species and

**Table 4-1: Contingency table confronting cpDNA haplotype lineages with Structure clusters among *I. batatas*, *I. trifida*, *Ipomoea* sp. and *I. tabascana*.**

Ancestry for nuclear SSRs, as inferred by STRUCTURE	cp DNA haplotype lineages				
	Southern <i>I. batatas</i>		Northern <i>I. batatas</i>		Diploid <i>I. trifida</i>
	<i>I. batatas</i>	<i>Ipomoea</i> . sp	<i>I. batatas</i>	<i>Ipomoea</i> . sp	<i>I. tabascana</i>
K1 (Southern <i>I. batatas</i> )	33		11		
K2 ( <i>I. trifida</i> )				7	65
K3 (Northern <i>I. batatas</i> )	19		16		
K4 (Northern <i>I. batatas</i> )	1	2	14	5	1
Admixed “ <i>I. trifida</i> / <i>I. batatas</i> ”		3	1	5	
Admixed “ <i>I. batatas</i> / <i>I. batatas</i> ”	17		21		1

“Admixed *I. trifida* / *I. batatas*” correspond to individuals sharing ancestry (< 80 %) for nuclear SSRs with K2 and any of the clusters K1, K3 or K4.

“Admixed *I. batatas* / *I. batatas*” share ancestry for nuclear SSRs between any two of the three *I. batatas* clusters K1, K3 or K4.



the low yield of storage roots in such crosses (Shiotani *et al.* 1990) may limit the selection of such hybrids by farmers.

#### Scenario 2: Multiple domestication events in space and time

The great divergence we have documented between the Northern and Southern gene pools could alternatively be the result of two independent events of domestication, either of a single widely distributed, geographically differentiated species, or of different *Ipomoea* species in each region. Numerous genetic, phylogenetic and cytogenetic studies have established that diploid *I. trifida* is the closest relative to *I. batatas*, but have failed to determine precisely if it was the only genome involved in the formation of the hexaploid sweet potato. In addition, the taxonomy of *I. trifida* is quite debated. Diploid *I. trifida* occurs in lowlands (sea level to 1000m) from Mexico to Caribbean South America (northern Colombia and Venezuela). The distribution of this species in the lowlands of north-western South America (southern Colombia, Ecuador, and Peru) is poorly documented; only three specimens from this region (collected from both sides of the Andes) have been referred to this species (<http://www.tropicos.org>). For over 50 years, there has been interest in tetraploid specimens discovered in Ecuador, Colombia, Guatemala and Mexico (Bohac *et al.* 1993). After taxonomic evaluation, it appeared that many of these are wild tetraploid *I. batatas* (Bohac *et al.* 1993). These results have raised the possibility that *I. batatas* also exists as a wild species, widely distributed from Mexico to Ecuador. If this is so, cultivated forms could have been domesticated directly from this wild tetraploid ancestor, perhaps through hybridization with other related wild species (Bohac *et al.* 1993).

If different *Ipomoea* species were involved in the formation of these two cultivated groups, we expected from our nuclear data to detect a multimodal distribution of allelic size range. Because nuclear microsatellite loci often evolve following a stepwise mutation model, alleles from divergent species are expected to differ in their size ranges. Problems of amplification and numerous null alleles would also not be surprising in this case. The relative homogeneity in distribution of allele size range for each group and the presence, in each group, of accessions with at least five alleles for each of 10 out of the 13 loci examined, indicate that more probably one species, widely distributed and exhibiting geographical differentiation from Central to South America, was domesticated. If more than one species was involved, these must have been very closely related.

Numerous cpSSR haplotypes were present within each group. In particular, within both Northern and Southern groups, we obtained more than one common and widely distributed haplotype. In interpreting gene genealogies, such haplotypes are considered ancestral (Posada and Crandall 2001). The presence of two such haplotypes is consistent with multiple domestications within an ancestrally polymorphic wild population, as has been proposed for *Phaseolus vulgaris* (Chacón *et al.* 2005). Alternatively, early in the process of diffusion, domesticated sweet potatoes may have captured chloroplasts from related wild populations in regions where the crop was introduced, as discussed above in treatment of scenario 1. Finally, mutations may have occurred post-domestication, although this would imply surprisingly rapid evolution at cpSSR loci.

Sweet potato is an ancient crop in Peru (Zhang *et al.* 2000; Ugent and Peterson 1988). The oldest remains of dried sweet potatoes, 8000 years old, originate from the caves of Chilca Canyon in Peru and roots have also been found in other more recent coastal Peruvian archaeological sites (Ugent and Peterson 1988). No ancient remains have been reported in Central American sites. These archaeological data confirm sweet potato as one of the oldest domesticates in the Americas, domestication having occurred more than 8000 years ago, and likely first in South America, as already suggested for common bean (Chacón *et al.* 2005).

While this second scenario appears more parsimonious than the first, the apparent lack of morphological divergence between Central American/Caribbean and South American sweet potato landraces raises questions. In crops known to have been domesticated two or more times independently from different wild pools, morphological differentiation is expected, because the different origins are usually associated with independent mutations favoured by selection (Allaby *et al.* 2008) that lead to different phenotypic traits. No subspecies, botanical varieties or varietal groups of sweet potato corresponding to the two gene pools we identified have so far been reported in the literature. On the other hand, there exists no comprehensive, statistically supported phenotypic analysis of New World sweet potato landraces. The morphological descriptors classically used to characterize sweet potato landraces are highly variable within each of the two groups. However, these

(a) Factorial analysis of the 24 cpSSR haplotypes. (b) Minimum spanning network of the 24 cpSSR haplotypes. One-bp mutation events are indicated by the name of the locus concerned. Circle size is proportional to the number of accessions that carry this haplotype. For a given haplotype, the proportion of accessions from Southern region and Northern region carrying this haplotype are represented respectively in grey and black.

descriptors concern traits that should be under diversifying human selection. Great variation within groups in traits such as leaf shape and the shape and colour of tuberous roots, as well as high overlap between the two groups, are thus not surprising. Stable characters used by taxonomists to differentiate species in *Ipomoea* series *Batatas* concern floral traits, in particular sepal and corolla shapes and sepal size (Bohac *et al.* 2003). Such traits are often absent from lists of morphological descriptors analyzed in sweet potato landraces, but they may offer the best chance to detect eventual differences between South American and Central American/Caribbean landraces.

Finally, when available molecular, archaeological, morphological and taxonomic evidence are all taken into account, we postulate that at least two domestication events occurred for sweet potato, one in the lowlands of north-western South America, and one in the lowlands of the Central America/Caribbean region. Sweet potato then diffused from these areas of origin throughout tropical America.

If this scenario is supported by further tests, sweet potato would join other American crop species for which evidence suggests multiple independent domestication events. Diamond (1997) postulated that geography may have made independent domestication events of the same plant particularly likely in the New World. According to his hypothesis, the demographic (and sometimes military) advantages conferred by agriculture allowed early farming peoples to expand their ranges, carrying their domesticated plants with them. This process would have been slower in the Americas, oriented north-south, than in Eurasia, oriented east-west: people and crops crossing latitudes encounter greater environmental barriers than do people and crops moving with longitude (Diamond 1997). There thus was more time for independent domestication in northern and southern centers. Taro (*Colocasia esculenta* L.) offers a possible example from the Old World. Based on AFLP results, it has been proposed that this plant was domesticated at least twice, once in South-east Asia and once in the Pacific (Kreike *et al.* 1999). Here, insular isolation may have contributed to facilitating independent domestication.

#### *Dispersal and diversification processes*

From both of its putative areas of origin, sweet potato diffused broadly within and among regions long before the arrival of Europeans, as attested to by the wide distribution of common and probably ancestral haplotypes throughout the area sampled.

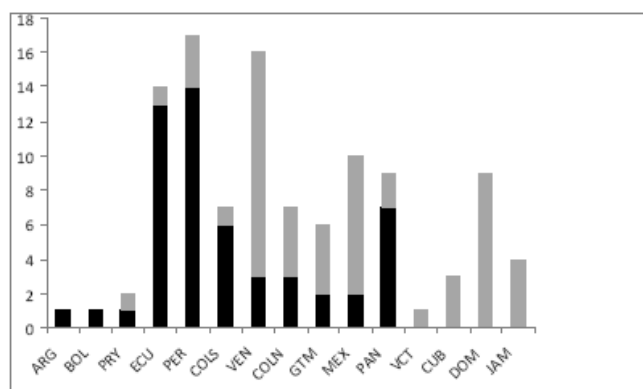
We identified six sets of clones with two or three replicates per set and also detected seven clonal lineages with two or three accessions per lineage. Most of the clone sets and clonal lineages were restricted to a single site or to sites in close geographic proximity. Comparing phylogeographic signals given by chloroplast and nuclear markers in the Northern region, we found very few accessions for which both types of markers indicated a Southern origin (i.e., group 1 haplotype, nuclear cluster 1). Similarly, in the Southern region, we found very few accessions for which both markers indicated a Northern origin (i.e., group 2 haplotype, nuclear cluster 2). In each region, most accessions that possessed “foreign” haplotypes belonged to the nuclear cluster typical of their respective areas. This indicates that recombination between migrant clones and local material likely occurs. Patterns of differentiation observed with nuclear data also indicate recombination between the two populations. Genetic differentiation at nuclear loci is less pronounced than that revealed by cpSSRs, while the phylogeographic signal still remains very strong. This pattern is probably due to admixture between the two initial domesticated gene pools. 87% of accessions displayed pairwise genetic distances high enough ( $>0.2$ ) to consider that they were likely to have been derived from independent recombination events rather than from somaclonal diversification (Arnaud-Haond *et al.* 2007). The incorporation of true seedlings issued from sexual reproduction has been reported in sweet potato landrace populations in Oceania (Yen, 1974), but its importance for diversification of American landraces has not been studied. Our data show that most of the genetic diversity we document cannot be accounted for by strictly clonal evolution, strongly suggesting the importance of mixed clonal/sexual reproduction via the incorporation of true seedlings as new clones, as in other clonally propagated domesticated plants (Elias *et al.* 2001; McKey *et al.* 2010). The seedlings farmers incorporated may in some cases have been wild/domesticated hybrids, accounting for the gene flow from wild to domesticated taxa suggested by our data.

The significant correlation between geographical and nuclear genetic distance observed within both the Southern and Northern regions probably reflects progressive expansion of the crop

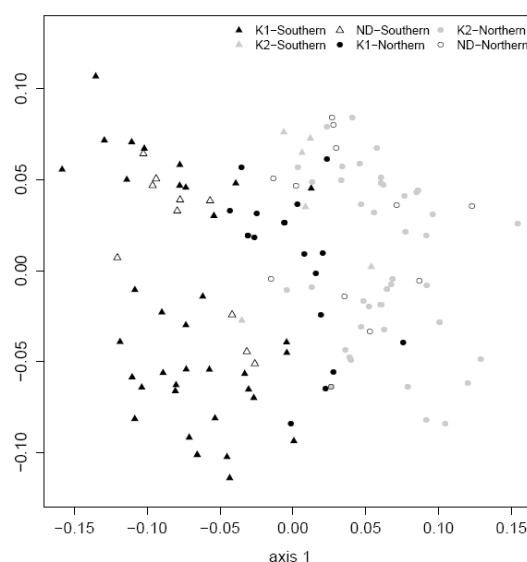
**Figure 3-3: Nuclear genetic diversity distribution of American sweet potato landraces.**

(a) Factorial analysis of the 130 accessions genotyped by nuclear SSRs. Accessions from the Southern region are represented by a triangle and those from Northern region by a circle. Colours indicate association with nuclear cluster 1 (K1; in black), cluster 2 (K2; in grey), missing data (ND; in white). (b) Census per country of accessions belonging to cluster 1 in black and cluster 2 in grey. Countries are labelled by their first three letters.

a)

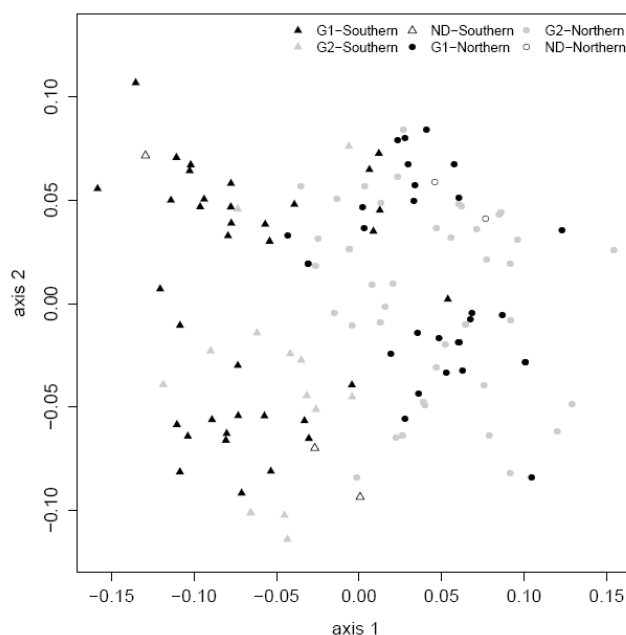


b)



**Figure 3-4: Factorial analysis of the 130 accessions genotyped with nuclear SSRs.**

Accessions from the Southern region are represented by a triangle and those from Northern region by a circle. Colours indicate association with chloroplast group 1 (G1; in black), chloroplast group 2 (G2; in grey), missing data (ND; in white). Accessions associated with cpSSR haplotype group 1 are in black, and accessions from group 2 are in grey.



within each region. Within the Southern region, there was greater sub-regional differentiation and a stronger pattern of isolation by distance than in the Northern region, possibly indicating greater barriers to gene flow within the Southern region.

In most clonally propagated crops, sexual reproduction has played a major role in evolution under domestication (Pickersgill 2007). The widespread notion of ‘single-step domestication’ of these crops by capture and multiplication of mutant clones appears to be untenable (McKey *et al.* 2010). Domestication probably occurred over a period of time, and perhaps in more than one location throughout the vast geographical ranges of most crop plants’ wild ancestors. In fact, the domestication process is still being practiced by farmers who capture attractive volunteer genotypes appearing spontaneously and multiply them as new clones. Different varieties of these crops exchange genes, and gene flow occurs between cultivated forms and wild taxa, or between cultivated forms and feral plants escaped from cultivation and surviving in the wild (Lebot 2009). Our results support such scenarios for sweet potato domestication and diversification in the Americas. We propose that sweet potato was probably domesticated at least two times, once in the lowlands of north-western South America, and once in the lowlands of Central America or the Caribbean rim. Domestication may even have occurred more than once within one or both regions. Sweet potato was then dispersed across tropical America. In the long run, social exchange of clones, somatic mutation, local recombination, and gene flow between wild and cultivated forms have contributed to shape the large number of landraces recorded in the Americas.

## ACKNOWLEDGEMENTS

This study was funded by grants from the “Pacific funds” distributed by the French Ministry of European and Foreign Affairs. C.R. was supported by a doctoral grant from the French Ministry of Higher Education and Research. The authors are very grateful to G. Talbot for his assistance on the "R" software, to Y. Vigouroux and M.I. Chacón for discussions about this work. We also wish to thank Marie-France Duval and CIP and CIRAD field helpers. All laboratory work was performed in the ‘Service Commun des Marqueurs Moléculaires en Ecologie’ of the CEFÉ, and sequences were obtained through the facilities of the IFR 119 ‘Montpellier Environnement Biodiversité’.

## REFERENCES

- Allaby RG, Fuller DQ, Brown TA (2008) The genetic expectations of a protracted model for the origins of domesticated crops. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 13982-13986.
- Arnaud-Haond S, Duarte CM, Alberto F, Serrão EA (2007) Standardizing methods to address clonality in population studies. *Molecular Ecology*, **16**, 5115-5139.
- Arrigo N, Felber F, Parisod C, Buerki S, Alvarez N, David J, Guadagnuolo R (2010) Origin and expansion of the allotetraploid *Aegilops geniculata*, a wild relative of wheat. *New Phytologist*, **187**, 1170-1180.
- Austin DF (1977) Hybrid polyploids in *Ipomoea* section *Batatas*. *Journal of Heredity*, **68**, 259-260.
- Austin DF (1987) The taxonomy, evolution and genetic diversity of sweet potatoes and related wild species. In: Exploration, maintenance, utilization of sweet potato genetic resources. Proceedings of the first Planning Conference, Lima, Peru, International Potato Center (CIP): 27-59.
- Bohac JR, Austin DF, Jones A (1993) Discovery of wild tetraploid sweet-potatoes. *Economic Botany*, **47**, 193-201.
- Bryan GJ, McNicoll J, Ramsay G, Meyer RC, De Jong WS (1999) Polymorphic simple sequence repeat markers in chloroplast genomes of solanaceous plants. *Theoretical and Applied Genetics*, **99**, 859-867.
- Burger JC, Chapman MA, Burke JM (2008) Molecular insights into the evolution of crop plants. *American Journal of Botany*, **95**, 113-122.
- Burnier J, Buerki S, Arrigo N, Kupfer P, Alvarez N (2009) Genetic structure and evolution of alpine polyploid complexes: *Ranunculus kuepferi* (Ranunculaceae) as a case study. *Molecular Ecology*, **18**, 3730-3744.
- Buteler MI, Jarret RL, LaBonte DR (1999) Sequence characterization of microsatellites in diploid and polyploid *Ipomoea*. *Theoretical and Applied Genetics*, **99**, 123-132.
- Caillon S, Quero-Garcia J, Lescure JP, Lebot V (2006) Nature of taro (*Colocasia esculenta* (L.) Schott) genetic diversity prevalent in a Pacific Ocean island, Vanua Lava, Vanuatu. *Genetic Resources and Crop Evolution*, **53**, 1273-1289.
- Chacón MI, Pickersgill B, Debouck DG (2005) Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. *Theoretical and Applied Genetics*, **110**, 432-444.

**Table 3-5: Nuclear diversity (number of alleles) of the Southern and Northern regions counting respectively 49 and 81 accessions.**

The value between brackets for allele number in Northern region corresponds to the mean number of alleles from 1000 re-samplings of 49 accessions.

Locus name	Allele number			Private alleles	
	Global	Southern region	Northern region	Southern region	Northern region
J315E	9	6	9 (8.61)		3
J522A	10	8	7 (6.51)	3	2
J263	7	7	5 (4.65)	2	
Ib297	23	12	21 (19.11)	2	11
J10A	10	8	9 (8.3)	1	2
J206A	9	9	5 (5)	4	
J116A	15	14	15 (14.42)		1
IbC5	12	9	11 (10.63)	1	3
IbR16	8	5	8 (7.8)		3
J130	4	4	4 (4)		
J1809E	8	6	7 (6.67)	1	2
IbS11	12	11	11 (10.56)	1	1
J544b	8	8	7 (6.96)	1	
total	135	107	119	16	28
<i>Mean</i>	<i>10.384</i>	<i>8.385</i>	<i>9.153 (8.654)</i>		

- Clement CR, De Cristo-Araújo M, Coppens d'Eeckenbrugge G, Alves Pereira A, Picanço- Rodrigues D (2010) Origin and domestication of native Amazonian crops. *Diversity*, **2**, 72-106.
- Diamond J (1997). Location, location, location: The first farmers. *Science*, **278**, 1243-1244.
- Dickau R, Ranere AJ, Cooke RG (2007) Starch grain evidence for the pre-ceramic dispersals of maize and root crops into tropical dry and humid forests of Panama. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 3651-3656.
- Elias M, Penet L, Vindry P, McKey D, Panaud O, Robert T (2001) Unmanaged sexual reproduction and the dynamics of genetic diversity of a vegetatively propagated crop plant, cassava (*Manihot esculenta* Crantz), in a traditional farming system. *Molecular Ecology*, **10**, 1895-1907.
- Excoffier, L. G. Laval, and S. Schneider (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47-50.
- Gepts P, Osborn TC, Rashka K, Bliss FA (1986) Phaseolin-protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris*) - evidence for multiple centers of domestication. *Economic Botany*, **40**, 451-468.
- Gichuki ST, Berenyi M, Zhang DP, Hermann M, Schmidt J, Glössl J, Burg K (2003) Genetic diversity in sweetpotato [*Ipomoea batatas* (L.) Lam.] in relationship to geographic sources as assessed with RAPD markers. *Genetic Resources and Crop Evolution*, **50**, 429-437.
- Hartigan JA, Wong MA (1979) A K-means clustering algorithm. *Applied Statistics*, **28**, 100-108.
- Hijmans RJL, Guarino L, Bussink C, Mathur P, Cruz C, Barrentes I, Rojas E (2004). DIVA-GIS Vsn. 5.0. A geographic information system for the analysis of species distribution data. Manual available at <http://www.diva-gis.org>.
- Huang JC, Sun M (2000) Genetic diversity and relationships of sweetpotato and its wild relatives in *Ipomoea* series *Batatas* (Convolvulaceae) as revealed by inter-simple sequence repeat (ISSR) and restriction analysis of chloroplast DNA. *Theoretical and Applied Genetics*, **100**, 1050-1060.
- Huang JC, Corke H, Sun M (2002) Highly polymorphic AFLP markers as a complementary tool to ITS sequences in assessing genetic diversity and phylogenetic relationships of sweetpotato (*Ipomoea batatas* (L.) Lam.) and its wild relatives. *Genetic Resources and Crop Evolution*, **49**, 541-550.
- Imazio S, Labra M, Grassi F, Scienza A, Failla O (2006) Chloroplast microsatellites to investigate the origin of grapevine. *Genetic Resources and Crop Evolution*, **53**, 1003-1011.
- Jarret RL, Austin DF (1994) Genetic diversity and systematic relationships in sweet potato (*Ipomoea batatas* (L.) Lam.) and related species as revealed by RAPD analysis. *Genetic Resources and Crop Evolution*, **41**, 165-173.
- Kobayashi M (1984) The *Ipomoea trifida* complex closely related to sweet potato. In: Shideler, S.F. and Rincon, H. (eds). Proceedings of the 6th Symposium of the International Society of Tropical Root Crops. CIP, Lima, Peru, pp. 561-568.
- Kreike CM, Van Eck HJ, Lebot V (2004) Genetic diversity of taro, *Colocasia esculenta* (L.) Schott, in Southeast Asia and the Pacific. *Theoretical and Applied Genetics*, **109**, 761-768.
- Lebot V (2009) Tropical roots and tubers crops: cassava, sweet potato, yams and aroids. Wallingford : CABI, XIX-413 pp.
- Lebot, V (2010) Sweet Potato. Chapter 3 In Bradshaw JE (ed.) Root and Tuber Crops. Handbook of Plant Breeding 7, 97-125. Springer, NY.
- Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA (2006) Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 9578-9583.
- Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez G J, Buckler E, Doebley J (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 6080-6084.
- McKey D, Elias M, Pujol B, Duputié A (2010) The evolutionary ecology of clonally propagated domesticated plants. *New Phytologist*, **186**, 318-332.
- Motta-Aldana JR, Serrano-Serrano ML, Hernandez-Torres J, Castillo-Villamizar G, Debouck DG, Chacón MI (2010) Multiple origins of lima bean landraces in the Americas: evidence from chloroplast and nuclear DNA polymorphisms. *Crop Science*, **50**, 1773-1787.
- Orjeda G, Freyre R, Iwanaga M (1991) Use of *Ipomoea trifida* germplasm for sweet-potato improvement. 3. Development of 4x interspecific hybrids between *Ipomoea batatas* (L.) Lam. (2n = 6x = 90) and *Ipomoea trifida* (Hbk) G Don (2n = 2x = 30) as storage root initiators for wild species. *Theoretical and Applied Genetics*, **83**, 159-163.
- Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288-295.



- Petit RJ, El Mousadik A, Pons O (1998) . Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844 – 855
- Perrrier X, Jacquemoud-Collet JP (2006) DARwin software <http://darwin.cirad.fr/darwin>
- Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution*, **16**, 37-45.
- Provan J, Powell W, Dewar H, Bryan G, Machray GC, Waugh R (1999) An extreme cytoplasmic bottleneck in the modern European cultivated potato (*Solanum tuberosum*) is not reflected in decreased levels of nuclear diversity. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **266**, 633-639.
- Provan J, Powell W, Hollingsworth PM (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology & Evolution*, **16**, 142-147.
- Pickersgill B (2007) Domestication of plants in the Americas: Insights from Mendelian and molecular genetics. *Annals of Botany*, **100**, 925-940.
- Rajapakse S, Nilmalgoda SD, Molnar M, Ballard, RE, Austin DF, Bohac JR (2004) Phylogenetic relationships of the sweetpotato in *Ipomoea* series *Batatas* (Convolvulaceae) based on nuclear beta-amylase gene sequences. *Molecular Phylogenetics and Evolution*, **30**, 623-632.
- Ranere AJ, Piperno DR, Holst I, Dickau R, Iriarte J (2009) The cultural and chronological context of early Holocene maize and squash domestication in the Central Balsas River Valley, Mexico. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5014-5018.
- Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp. 365-386.
- Sanjur OI, Piperno DR, Andres TC, Wessel-Beaver L (2002) Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from a mitochondrial gene: Implications for crop plant evolution and areas of origin. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 535-540.
- Shiotani I, Yoshida S, Kawase T (1990) Numerical taxonomic analysis and crossability of diploid *Ipomoea* species related to the sweet potato. *Japanese Journal of Breeding*, **40**, 159-174.
- Spooner DM, McLean K, Ramsay G, Waugh R, Bryan GJ (2005) A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 14694-14699.
- Srisuwan S, Sihachakr D, Siljak-Yakovlev S (2006) The origin and evolution of sweet potato (*Ipomoea batatas* Lam.) and its wild relatives through the cytogenetic approaches. *Plant Science*, **171**, 424-433.
- Ugent D, Peterson W (1988) Archeological remains of potato and sweet potato in Peru. *CIP Circular*, **16**, 1-10.
- Vigouroux Y, Glaubitz JC, Matsuoka Y, Goodman MM, Jesus SG, Doebley J (2008) Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. *American Journal of Botany*, **95**, 1240-1253.
- Van Heerwaarden J, Doebley J, Briggs HJ, Glaubitz JC, Goodman MM, Sanchez Gonzalez JJ, Ross-Ibarra J (2011) Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 1088-1092
- Wills DM, Burke JM (2006) Chloroplast DNA variation confirms a single origin of domesticated sunflower (*Helianthus annuus* L.). *Journal of Heredity*, **97**, 403-408.
- Xu DH, Abe J, Gai JY, Shimamoto Y (2002) Diversity of chloroplast DNA SSRs in wild and cultivated soybeans: evidence for multiple origins of cultivated soybean. *Theoretical and Applied Genetics*, **105**, 645-653.
- Yen, DE (1974) *The Sweet Potato in Oceania: An Essay in Ethnobotany*. Bernice P. Bishop Museum 236, Honolulu, Hawaii: Bishop Museum Press.
- Zhang DP, Cervantes J, Huaman Z, Carey E, Ghislain M (2000) Assessing genetic diversity of sweet potato (*Ipomoea batatas* (L.) Lam.) cultivars from tropical America using AFLP. *Genetic Resources and Crop Evolution*, **47**, 659-665.

## DATA ACCESSIBILITY

**Complete sample passport data and chloroplast / nuclear microsatellites data (alleles size tables):** Data deposited at Dryad: doi:10.5061/dryad.5h5q4

# CHAPITRE 4

*Du sauvage au domestique*

3) *Amérique tropicale*

*Article : Relations phylogénétiques entre Ipomoea batatas et les espèces sauvages apparentées*

*Disentangling the origins of cultivated sweet potato  
(Ipomoea batatas (L.) Lam.)*

- 1. Title:** Disentangling the origins of cultivated sweet potato (*Ipomoea batatas* (L.) Lam.))
- 2. Authors' names:** C. Roullier<sup>1,2,3</sup>, A. Duputié<sup>2</sup>, P. Wennekes<sup>2</sup>, L. Benoit<sup>2</sup>, V. M. F. Bringas<sup>4</sup>, G. Rossel<sup>4</sup>, D. Tay<sup>4</sup>, D. McKey<sup>2,3</sup>, V. Lebot<sup>1</sup>
- 3. Full postal address:**  
<sup>1</sup> CIRAD, Avenue Agropolis, 34398 Montpellier cedex 5, France  
<sup>2</sup> Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175, 1919 route de Mende, 34293 Montpellier cedex 5, France  
<sup>3</sup> Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier cedex 5, France.  
<sup>4</sup> International Potato Center (CIP), Lima, Peru
- 4. Keywords:** *Ipomoea batatas*, polyploidy, domestication, chloroplast haplotypes, nuclear microsatellites, phylogeography
- 6. Corresponding author:** Roullier Caroline, CIRAD Avenue Agropolis, 34398 Montpellier Cédex 5 France; E-mail address: [carolineroullier@hotmail.com](mailto:carolineroullier@hotmail.com).
- 6. Running title:** Polyploidization history in sweet potato /
- 7. Article submitted to Plos ONE. Accepted with minor revisions.**

## ABSTRACT

Sweet potato (*Ipomoea batatas* (L.) Lam., Convolvulaceae) counts among the most widely cultivated staple crops worldwide, yet the origins of its domestication remain unclear. It is assumed that this hexaploid species could have had either an autopolyploid origin, from the diploid *I. trifida*, or an allopolyploid origin, including genomes of *I. trifida* and *I. triloba*. We typed a broad sample of cultivated sweet potatoes and its diploid and polyploid wild relatives for noncoding chloroplast and nuclear sequences, and nuclear SSRs. *Ipomoea trifida* and *I. batatas* form a monophyletic clade. The distant relationship between *I. batatas* and *I. triloba*, shown both by chloroplast and nuclear markers and the overall similar allelic composition of *I. trifida* and *I. batatas* for nuclear SSRs, support an autopolyploid origin of sweet potato from the ancestor it shares with *I. trifida*, which might be similar to currently observed tetraploid wild *Ipomoea* accessions. Two *I. batatas* chloroplast lineages were identified; these exhibit more divergence between each other than each does with *I. trifida*. We thus propose that cultivated *I. batatas* have multiple origins, and evolved from at least two distinct autopolyploidization events in polymorphic wild populations of a single progenitor species. Secondary contact between sweet potatoes domesticated in Central America and in Southern America, from differentiated wild *I. batatas* populations, would have led to the introgression of chloroplast haplotypes of each lineage into nuclear backgrounds of the other, and to a reduced divergence between nuclear gene pools as compared with chloroplast haplotypes.

## INTRODUCTION

Polyploidy is recognized as an important factor in the evolution and diversification of plants [1]. Polyploid crops are common, and include for example banana, bread wheat, potato, sugar beet and sweet potato, and polyploidy is frequently used by breeders for crop improvement. Crop domestication corresponds to an evolutionary process of species divergence, in which genetic, morphological and physiological changes result from the cultivation of plants by humans [2]. Often considered an “event”, particularly for clonally propagated crops [3], domestication is increasingly looked upon as a protracted process, involving repeated recombination-selection cycles and often wild/cultivated gene flow, with artificial (conscious or not) and natural selection interacting to drive the wild-to-cultivated transition [4,5]. The link between polyploidy and domestication is not clearly established, although some have speculated that polyploidy may predispose crops for domestication [6].

In natural populations, polyploid species may be formed through several mechanisms. Classically, autopolyploidy (genome duplication with a single progenitor species) has been distinguished from allopolyploidy (hybridization and genome doubling of highly divergent parental species; [1]). However, there is a continuum between the two. Autopolyploid complexes often have multiple independent origins, sometimes involving crosses between conspecific, but still substantially differentiated, populations [7]. Polyploidization often triggers genomic re-patterning and gene expression changes [1], which could explain the sudden appearance of new phenotypes that diverge from those of their diploid parents in numerous traits. Although these genetic changes are probably more rapid and extensive in allopolyploids, they may also affect autopolyploids over the longer term [7]. Moreover, polyploids may be reproductively isolated from their parents, and often can adapt to new ecological niches [1]. Shifts towards higher ploidy levels thus often drive speciation in plants [8], and indeed appear as a clear route to sympatric speciation [9]. In this context, autopolyploidy seems to have a higher incidence than previously assumed [7]. While the proportion of polyploids among crops is not statistically different from that among wild species of the same families [10], in some cases, polyploidy certainly provided raw material to achieve plant domestication. For example, the exploitation of fertile diploid banana genotypes began in New Guinea in the early Holocene. Human-mediated transfers of these diploids through islands of Melanesia and South-Eastern Asia have allowed hybridizations between allopatric subspecies of diploid *Musa acuminata*. As an evolutionary consequence of this hybrid status, seedless triploid genotypes were formed and widely cultivated [11].

Numerous crops were domesticated in the Americas, including potato, tomato, manioc, maize, beans, sweet potato, squash and many others [12]. Even though some crops were domesticated only once, in a restricted area (as is the case for maize; [13,14]), most of them were domesticated over a diffuse area [15], and some were domesticated two or more times independently. Common bean [16] and Lima bean [17] were domesticated several times each, in different parts of the range of their respective ancestral species, and the domesticated gene pools came into contact only secondarily. Whether a particular crop was domesticated once or several times, and from which species, is of more than historical interest. Independent domestication events raise the fundamental issue of determining how multiple genetic paths could lead to similar domestication traits, and documenting independent domestication provides essential data for the management of crop diversity, for the conservation of genetic resources, and for the use of wild relatives in breeding programs.

Sweet potato (*Ipomoea batatas* (L.) Lam., Convolvulaceae) is a clonally propagated hexaploid crop native to the Americas. It is a major staple crop, particularly in numerous developing countries [18]. Despite its importance, the botanical origin of sweet potato and the timing and geographic location(s) of its domestication remain unclear. Plants classified as *Ipomoea batatas* are mostly cultivated, hexaploid clones. Several molecular-genetic studies appeared to indicate that the diploid *I. trifida* is the closest wild relative of cultivated sweet potato [19,20,21,22,23,24], thus pointing to this species as the most likely candidate progenitor of sweet potato. Two hypotheses have been proposed as to the origin of cultivated sweet potato. On the one hand, Austin [25] proposed that natural hybridization between *I. trifida* and *I. triloba* could have generated the ancestors of sweet potato somewhere between the Yucatan peninsula and the Orinoco basin. On the other hand, Kobayashi [26]

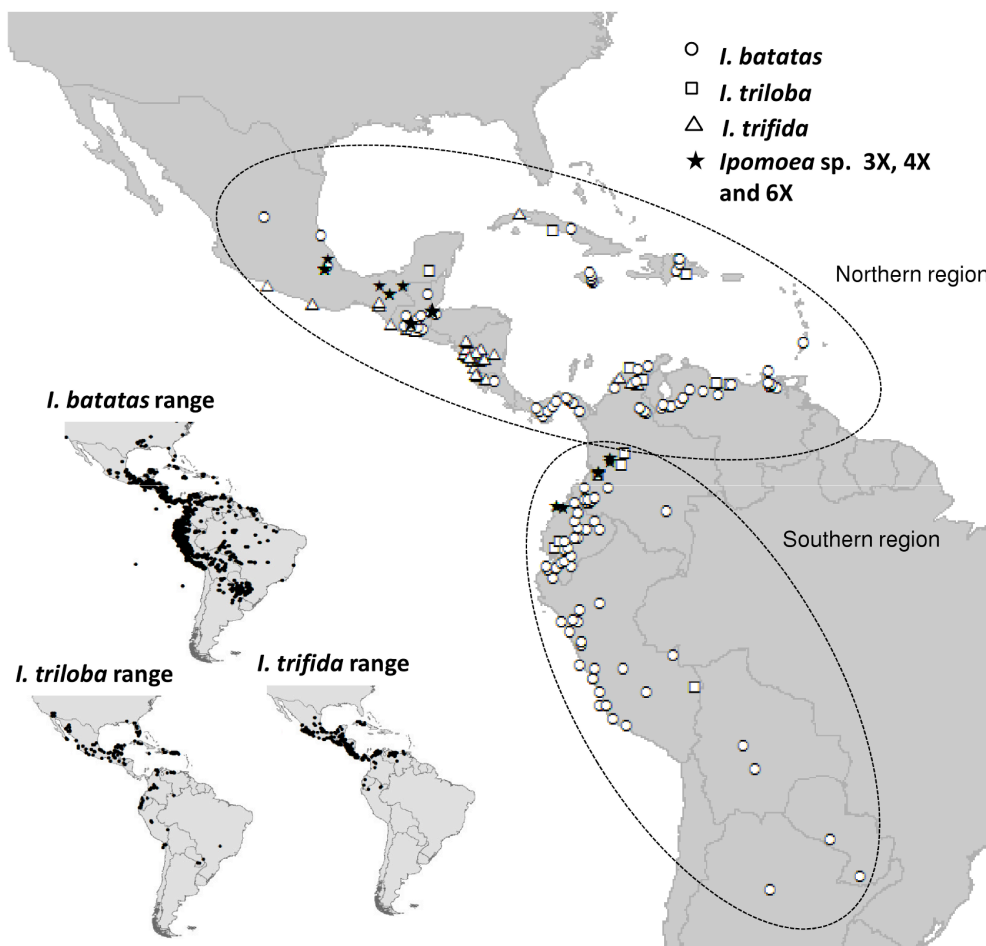
proposed that *I. trifida* forms an autopolyploid complex, with ploidy levels ranging from diploid to hexaploid, and that domesticated sweet potato derives from this group.

However, several accessions of *Ipomoea* with various ploidy levels (mostly 4X, but also a few 3X and 6X) are clearly closely related to *I. batatas*, but their taxonomic status is disputed and they are poorly characterized genetically. Collected in Ecuador, Colombia, Guatemala and Mexico, these accessions were initially identified as *I. trifida*, and after re-examination, most of them were re-identified as wild *I. batatas* on taxonomical grounds [27]. Whether these could represent wild *I. batatas*, or feral forms, has not been investigated. However, their discovery strongly suggested the possibility that *I. batatas* includes not only the hexaploid cultigen but also true wild populations with lower ploidy levels, from which cultivated forms would have been domesticated. Whether domestication preceded polyploidization (as in bananas, [11]) or followed it (as in cotton, [28]) is still unknown. Furthermore, recent findings point towards a possible double origin of domestication for sweet potato: cultivated landraces comprise two distinct geographically clustered chloroplast lineages [29], one corresponding mostly to landraces cultivated in Central America and the Caribbean (hereafter termed the Northern lineage), and the other to landraces found in northwestern South America (hereafter Southern lineage). Nuclear microsatellite markers showed similar differentiation patterns, supporting the idea that this divergence may predate sweet potato domestication [29]. The Northern and Southern genepools could thus represent two independent polyploidization and domestication histories.

In the present study, we address the issues of the botanical and geographic origin of sweet potato, using a representative sampling of *I. batatas*, its putative close wild diploid relatives *I. trifida* and *I. triloba*, and polyploid *Ipomoea* sp., using several neutral nuclear and chloroplast markers. We pose the following questions: Are *I. batatas* gene pools derived from a single progenitor, from differentiated conspecific populations that hybridized, or from multiple hybridizing species? Can we identify the progenitor(s) and pinpoint the geographical origin(s) of domestication? The formation of sweet potato's hexaploid genome must have involved at least two steps, from diploidy to intermediate ploidy levels (triploid or tetraploid) and then hexaploidy. Can we establish the temporal sequence of domestication relative to polyploidization?

**Figure 4-1: Geographical distribution of the sampling.**

Location of *I. triloba*, *I. trifida*, *I. batatas* and polyploid *Ipomoea* sp. accessions used in the present study and current taxa distribution ranges, as determined from GBIF records (<http://data.gbif.org/species/>) are provided. Accessions with no geographical information are not shown; details on sampling are provided in table S1.



## MATERIALS AND METHODS

### *Sampling*

Overall, 297 leaf tissue samples were collected on 216 accessions, representing six species from *Ipomoea* series *Batatas* as well as samples of polyploid accessions of dubious taxonomy. Accessions are synonymous to clones for the cultivated *I. batatas* and to populations for the wild relatives. Sampling for wild taxa includes 1-11 (median 2) individuals per population.

*Ipomoea batatas* (139 accessions) was sampled throughout its distributional range from Mexico to Peru, to represent both domesticated gene pools ([29]; 139 accessions). *Ipomoea trifida* (37 accessions) and *I. triloba* (15 accessions) were sampled throughout their distributional ranges (Figure 4-1). One accession of *I. tabascanana* was included, since this taxon has been postulated to be a hybrid between *I. batatas* and *I. trifida* [30]. Three accessions of *I. leucantha* and four of *I. tiliacea* were sampled as outgroups. Samples of dubious taxonomical status (hereafter *Ipomoea* sp.), sometimes referred to as wild and/or feral *I. batatas* [27] and sometimes as *I. trifida* [26, 31], were included (15 accessions; one diploid, one triploid, 11 tetraploid and two hexaploid accessions). Accessions were collected from the International Potato Center (CIP, Lima, Peru), the Genetic Resources Program of the Germplasm Resources Information Network (GRIN, USDA, USA), and the Japanese National Institute for Agrobiological Sciences (NIAS, Tsukuba, Japan; Table S1). Plants were grown from seeds or tubers in greenhouses. Leaf tissue was collected and dried and DNA was extracted using the Qiagen 96 Plant kit for lyophilised tissues (Hilden, Germany). For some accessions, ploidy level was verified by flow cytometry.

### *DNA sequencing*

The chloroplast intergenic spacer (IGS) *rpl32F-trnL* [32] were sequenced for all wild accessions and for a subsample of 23 cultivated *I. batatas* representing both cp lineages as defined in Roullier et al. [29] (164 samples in total). The nuclear ITS region (ITS4-ITS5; [33]) was sequenced only for a subsample of individuals (95 wild and cultivated samples). PCRs were performed in a final volume of 20  $\mu$ L, using 2  $\mu$ L of 1:10 diluted template DNA, 20  $\mu$ M of each primer, and 10  $\mu$ L of Phusion Taq mastermix (New England Biolabs, Inc.). Amplification was performed after 5 minutes denaturation at 98°C, over 30 cycles of 30 s denaturation at 98°C, 1 minute annealing at 57°C and 1 minute elongation at 72°C, and a final elongation of 5 minutes, using a PTC-100 Thermocycler (MJ Research, Waltham, MA, USA). Fluorescent dye-terminator sequencing was performed by Agowa GmbH (Berlin, Germany). All DNA samples were sequenced in forward and backward directions, and ten random duplicates were sequenced for quality control (all duplicates gave congruent results).

### *Sequence analysis and phylogenetic reconstruction*

Forward and reverse chromatograms were assembled and eye-checked independently by two investigators using Sequencher (Gene Codes, Ann Arbor, MI, USA). ITS sequences resulted in numerous heterozygotes. Length-variant heterozygotes were systematically discarded, while length-invariant heterozygotes were retained and coded with ambiguous character states. Only those chromatograms that produced a clear consensus were used in the following analyses (164 IGS sequences and 66 ITS sequences, in total). All sequences were deposited in GenBank. Sequences were aligned using Muscle (Edgar 2004, using the EMBL web service available at <http://www.ebi.ac.uk/Tools/msa/muscle/>). The resulting alignment was edited using BioEdit (Hall 1999). All mononucleotide repeats were discarded since these are prone to homoplasy [34], and indels were coded as binary characters, using the simple method of Simmons & Ochoterena [35].

Bayesian reconstruction of the phylogenetic tree was performed for IGS using MRBAYES [36], using the *I. purpurea* NCBI GenBank database sequence AY53831 as an outgroup. The most likely model for sequence evolution was selected among those implemented in FINDMODEL (<http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html>) using the Akaike Information Criterion. Since FINDMODEL does not admit binary character coding, model selection did not take indels into account. MRBAYES was run four times with four chains, for 10 million iterations. Convergence was attained after 2.5 million iterations, which were discarded as burn-in. A statistical



parsimony haplotype network was also constructed for IGS using TCS 1.21 [37], with indels treated as a fifth character state.

Using the program SPLITSTREE version 4.6 [38], we constructed an unrooted neighbour-joining tree based on Hamming distance (which handles ambiguous characters) for ITS sequences, whose robustness was assessed through 1,000 bootstrap resamplings.

### *SSR genotyping*

283 individuals (belonging to 7 taxa) were genotyped for eight nuclear microsatellites (J263, J522A, Ib297, J206A, IbR16, IbC5, J544b, IbS11) described in a previous study [29]. All loci were amplified independently using Multiplex PCR Taq (Qiagen) in a final volume of 10  $\mu$ L, using 30ng of DNA per reaction. The following programme was conducted using a PTC-100 Thermocycler (MJ Research, Waltham, MA, USA): 15 min at 95°C, 35 cycles of 30 s at 94°C, 1 min 30 s at 57°C, 1 min at 72°C and finally 30 min at 72°C. Allele scoring was eye-checked by two investigators using GENEMAPPER (Applied Biosystems, Foster City, CA, USA).

### *Taxon boundaries as assessed from SSR data*

We first represented the genetic relationships between *I. batatas*, its five wild relatives and *Ipomoea* sp. accessions performing a Principal Coordinate Analysis (PCoA) based on a Lynch distance matrix between individuals, using R software (R Development Core Team 2011) and the polysat package [41]. We also constructed an unrooted neighbour-joining tree based on the same dissimilarity measure with the ape package [39] and edited it using Darwin 5 software [40].

Genetic relationships between the closely related *I. trifida*, *I. batatas* and *Ipomoea* sp. were further examined by clustering methods. First, model-based Bayesian clustering was run using the software STRUCTURE 2 [42,43], assuming an admixture model, with allelic frequencies correlated among clusters, and dominant markers coding. 1.5 million MCMC steps were performed, with the first 500,000 iterations discarded as burn-in. To determine the optimal number of clusters  $K$ , we used the criterion derived by Evanno *et al.* [44], after having run STRUCTURE 8 times for  $K$  from 2 to 8.

As it is unclear whether model-based clustering methods are appropriate for treating data on dominant markers in a context of polyploidy, especially when data on microsatellite alleles must be converted to binary data, we compared results of this analysis with those from non model-based clustering approaches. Such approaches have been successfully used to detect differentiated polyploid lineages [45,46], and to study gene flow between polyploid taxa [47]. We first performed a non-hierarchical K-means clustering. This technique assigns individuals to  $K$  genetic groups, so as to maximize intergroup variance (quantified by inertia). We performed 50,000 independent clusterings for  $K = 2$  to  $K = 8$  using R scripts developed by Arrigo *et al.* [45], and determined the most likely number of clusters, as proposed by these authors. In order to determine whether non model-based clustering methods reproduced the results found by model-based clustering, we finally performed a fuzzy c-means clustering analysis using the method implemented in the 'e1071' R CRAN package [47]. This algorithm allocates individuals to a predefined number of groups so as to minimize intragroup variance, and calculates for each individual a probability of membership in each group. This membership coefficient can be compared to ancestry values inferred by STRUCTURE, and its computation relies on the fuzzification parameter  $r$  (ranging between one and infinity) determining the stringency of clustering: the larger the value of  $r$ , the fuzzier the clustering. We used  $K = 4$  clusters, since this was the most likely number of clusters found in the STRUCTURE and the K-means clustering, and tested values for  $r$  ranging from 1.1 to 1.5. Only low values of  $r$  gave relevant grouping results, and we present results for  $r = 1.1$ .

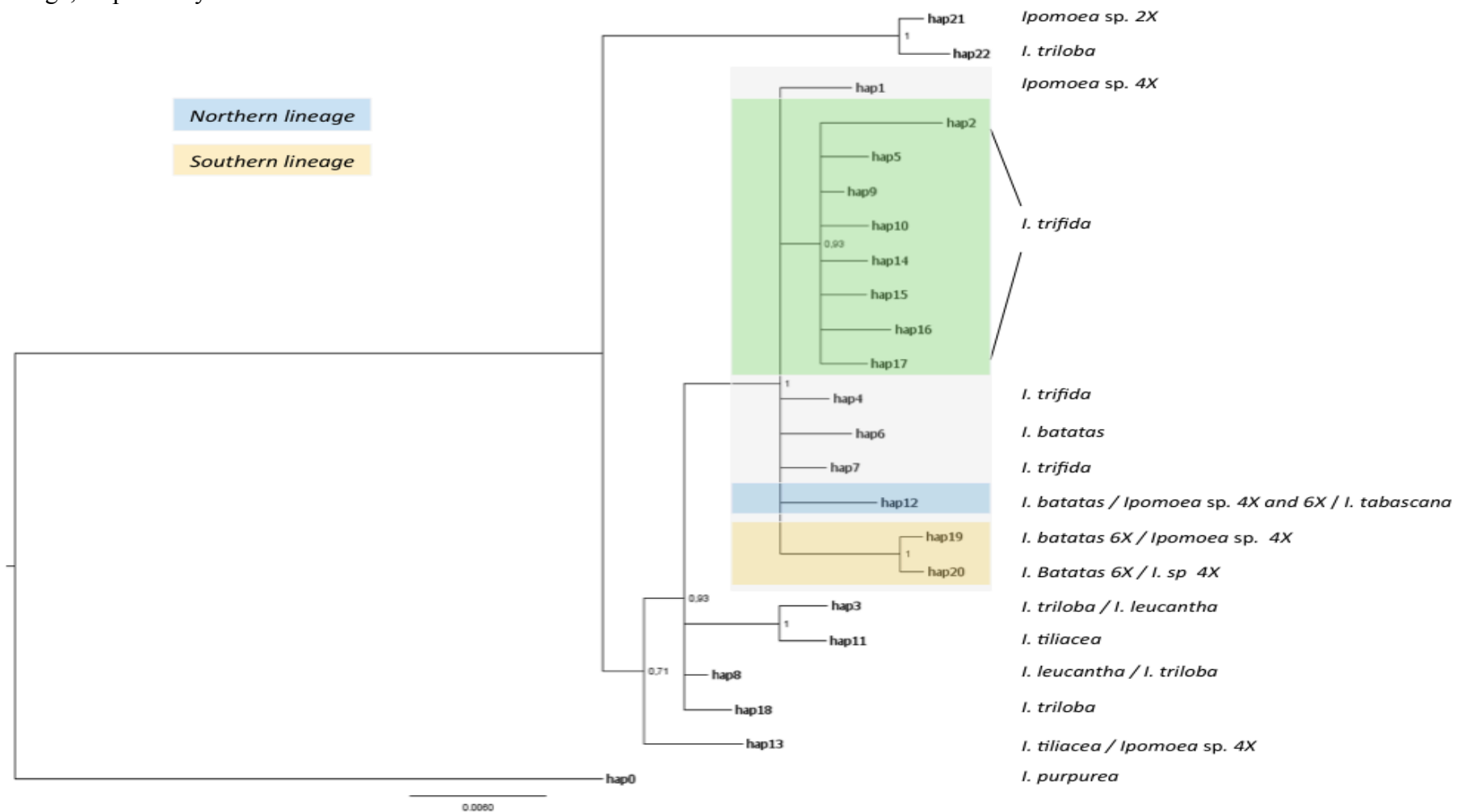
### *SSR diversity analysis*

We characterized the genetic diversity present in geographically well-sampled species (*I. batatas*, *I. trifida*, *I. triloba* and *Ipomoea* sp.), by computing the mean number of alleles per locus ( $NA$ ), its rarefied value  $Ar$  (averaged from 1000 resamplings of 25 individuals), the observed and expected heterozygosities  $Ho$  and  $He$  (the latter was determined for polyploid taxa from allelic frequencies

estimated using the R package polysat; “simpleFreq” method), and the intra-taxon mean pairwise Lynch distance between genotypes  $D$ .

**Figure 4-2: Genetic relationships of *I. batatas*, five wild relatives and *Ipomoea* sp. accessions based on chloroplast DNA analyses.**

a) *IGS* haplotypes majority rule consensus tree. Numbers along branches indicate bayesian posterior probabilities. Taxonomic identification associated with each haplotype is provided. Diploid *I. trifida*, Northern *I. batatas* and Southern *I. batatas* lineages are shown in green, blue and orange, respectively.



## RESULTS

### *Interspecific relationships as inferred from cpDNA sequences*

The 1077-bp long alignment of 164 IGS sequences showed 65 polymorphic sites, 19 of which were parsimony-informative, and 14 indels (once mononucleotide repeats were removed) resulting in 23 haplotypes.

Despite extensive geographic sampling of *I. trifida*, *I. triloba* and *I. batatas*, we found no haplotypes shared between any two of these species. *Ipomoea batatas*, *I. trifida* and *I. tabascanana* form a consistent monophyletic group (Bayesian posterior probability of 1; Figure 4-2a), together with some *Ipomoea* sp. samples, but excluding any *I. triloba*.

Out of 72 samples, 61 *I. trifida* shared haplotype 9 and the others carried haplotypes derived from this haplotype by one or two mutation steps (Figure 4-2b). Only four haplotypes were found over the 139 samples of *I. batatas*. As found by Roullier et al. [29], two distinct chloroplast lineages were found in *I. batatas*, mostly corresponding to Northern and Southern accessions (Figure 4-2b). They were more distantly related to each other than each was to *I. trifida*.

The *I. tabascanana* sample and numerous samples of dubious taxonomy (triploid, tetraploid and hexaploid *Ipomoea* sp.) carried the typical Northern *batatas* haplotype, while five tetraploid *Ipomoea* sp. samples carried a Southern *batatas* haplotype (three of them originated from Ecuador and two from Mexico). The unique diploid *Ipomoea* sp. carried a haplotype very close to that borne by one accession labelled as *I. triloba*, but distantly related to other *I. triloba* haplotypes, suggesting they may together form a distinct species. Finally, one tetraploid *Ipomoea* sp. sample bore a haplotype specific to *I. tiliacea*.

Concerning other species, phylogenetic relationships are less clearly resolved (Figure 4-2a). Moreover, some haplotypes are shared by accessions identified as different species, suggesting misidentifications (for example, haplotype 3 is shared among three species, *I. triloba*, *I. leucantha* and *I. tiliacea*).

### *Interspecific relationships as inferred from nuclear markers*

For 66 ITS sequences in total, 43 haplotypes were obtained; aligned sequences were 644 bp long. Consistent with the findings on cpDNA sequences, *I. batatas* shared no ITS sequences with *I. trifida* nor with *I. triloba*. The unrooted neighbour-joining tree for ITS sequences based on Hamming distance showed that haplotypes were mostly grouped by species (Figure 4-3a). *Ipomoea batatas* haplotypes formed two distinct groups, possibly corresponding to the two chloroplast lineages.

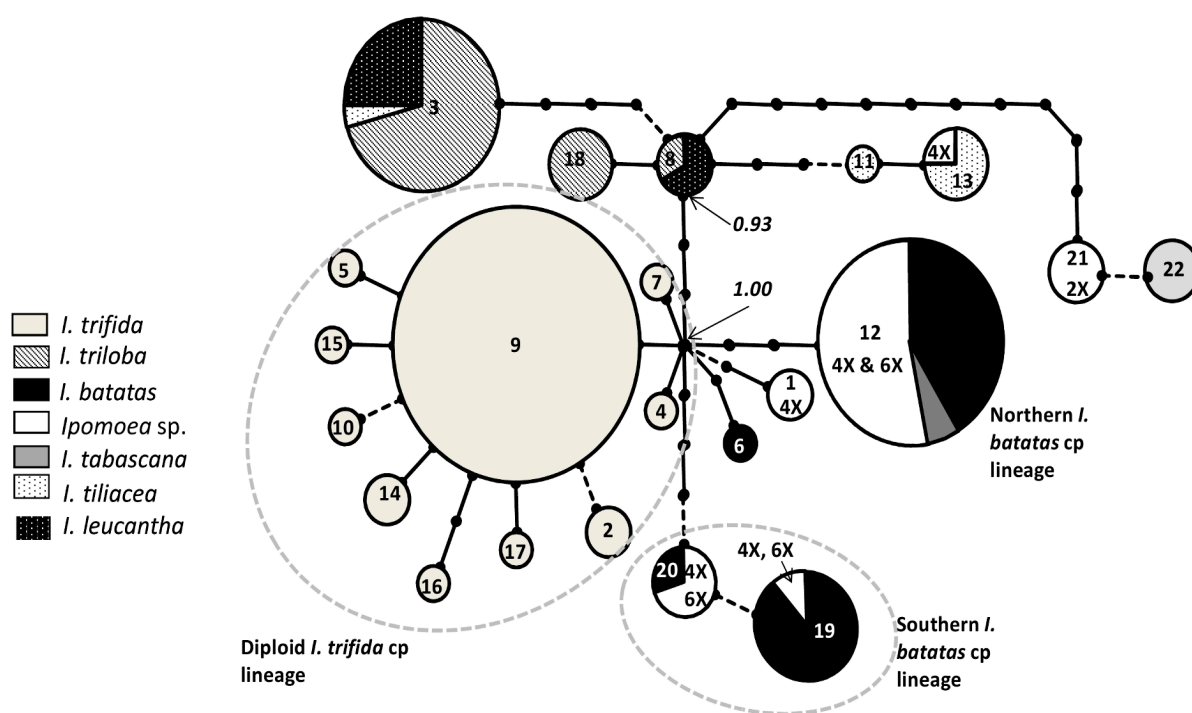
SSRs could be amplified for all loci and all species, leading to a total of 137 alleles. The unrooted neighbour-joining tree (Figure 4-3b) and the PCoA (Figure 4-3c) of SSR data showed consistent results. *Ipomoea trifida*, *I. triloba* and *I. batatas* formed three clear specific groups. While *I. tabascanana* and some *Ipomoea* sp. accessions were intermediate between *I. batatas* and *I. trifida* in both analyses, other *Ipomoea* sp. accessions were nested within *I. trifida*.

### *Genetic structure of I. trifida, I. batatas and Ipomoea sp.*

For the model-based clustering analysis, the appropriate number of clusters was unclear (Supplementary material 1). We thus present results for the values of  $K$  that provided the least admixed genotypes and are most relevant to describe our dataset:  $K = 2$  and  $K = 4$ . For  $K = 2$ , two distinct clusters grouped *I. batatas* accessions on the one hand, and *I. trifida* on the other. *Ipomoea* sp. accessions were either assigned (ancestry > 0.8) to one of the two groups, or shared ancestry within each group (not shown). For  $K = 4$ , *I. trifida* formed a well-defined cluster K2, to which all *I. trifida* accessions were assigned with more than 80 % ancestry (Figure 4-4a). Some tetraploid *Ipomoea* sp. accessions from Colombia were also clearly assigned to this cluster. Cluster K1 grouped most *batatas* accessions from the Southern region. *Ipomoea batatas* accessions from the Northern region were subdivided in two clusters: K3 and K4, with K3 grouping most of them, and K4 grouping 16

**Figure 4-2: Genetic relationships of *I. batatas*, five wild relatives and *Ipomoea* sp. accessions based on chloroplast DNA analyses.**

b) Statistical Parsimony Network of IGS haplotypes for *I. batatas*, five wild relatives and *Ipomoea* sp. accessions. Circle size is proportional to the number of individuals per haplotype. Substitutions and inversions are represented using full lines and indels are displayed using broken lines. Intermediate, unsampled haplotypes appear as dots. The posterior probability of two nodes, as obtained through a Bayesian tree reconstruction (Figure S4-1), is reported in italics. The ploidy level of *Ipomoea* sp. accessions is indicated.



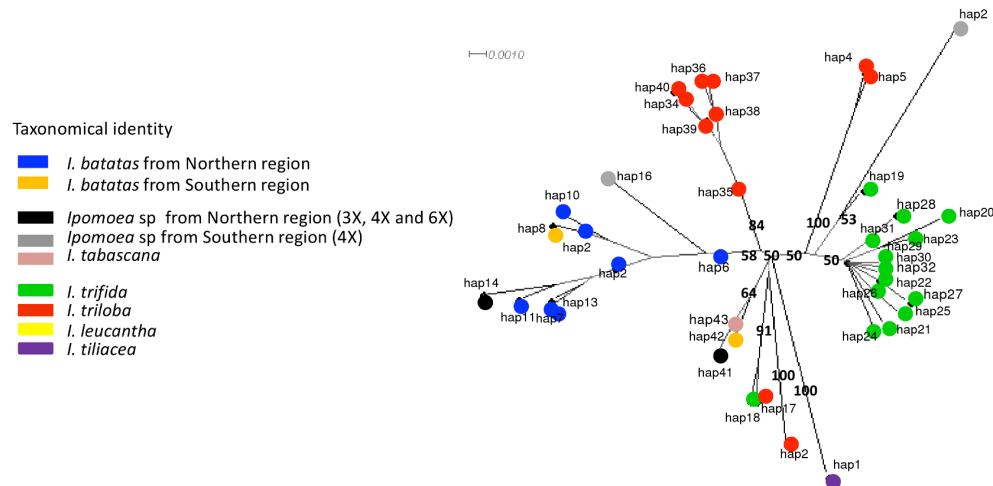
cultivated *I. batatas* from different countries in the Northern region, together with tetraploid and hexaploid *Ipomoea* sp. accessions, mostly from Mexico. With  $K = 4$ , 39 cultivated *I. batatas* (from both regions) had mixed ancestry (38 between the two *I. batatas* clusters and one between *I. batatas* and *I. trifida* clusters). This was also the case for 10 *Ipomoea* sp. accessions from Ecuador, Guatemala and Mexico, which had mixed ancestry between *I. batatas* clusters or between *I. trifida* and *I. batatas* (Figure 4a, Table 4-1).

For  $K = 4$ , fuzzy c-means clustering yielded results congruent with those obtained using the STRUCTURE approach. However, the proportion of “admixed” individuals (probability of membership to any cluster  $< 80\%$ ) was higher (Supplementary material 2). For K-means clustering,  $K = 2$  and  $K = 4$  were relevant groupings to describe our data (Supplementary material 2). K-means clusters were again consistent with the model-based clustering (Figure 4-4b). In this analysis, however, the 50 individuals determined as being of mixed ancestry in the STRUCTURE analyses were obligatorily assigned to one of the groups. For  $K = 4$ , 21 *I. batatas* were thus assigned to the Southern *batatas* cluster K1 (six of them were actually from South America); two *I. batatas* (from Colombia and Guatemala) and four *Ipomoea* sp. accessions (from Ecuador and Mexico) were assigned to the *trifida* cluster K2; 14 *I. batatas* (12 of which actually came from the Northern region) were grouped with the Northern *batatas* cluster K3; and six *I. batatas* (four of which came from the Northern region) and three tetraploid *Ipomoea* sp. accessions were assigned to the other Northern *batatas* cluster K4.

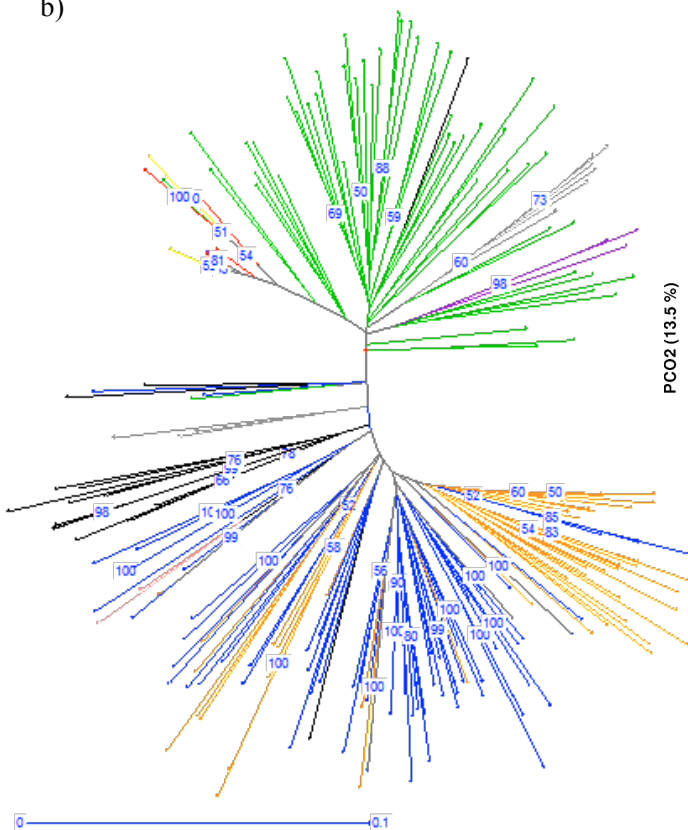
**Figure 4-3: Genetic relationships of *I. batatas*, five wild relatives and *Ipomoea* sp. accessions based on nuclear DNA analyses.**

a) Unrooted neighbour-joining tree based on Hamming distance between ITS haplotypes. Bootstrap values are indicated for central nodes. b) Unrooted neighbour-joining tree based on the Lynch distance between multilocus nuclear SSR genotypes c) Principal Coordinate Analysis based on the Lynch distance between multilocus nuclear SSR genotypes.

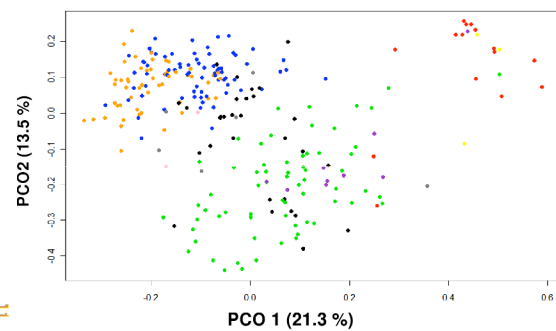
a)



b)



c)





### *Congruence between cpDNA haplotype groups and nuclear SSR genetic structure*

Both kinds of markers identified diploid *I. trifida* as a uniform genetic group (Figure 4-5). We did not sequence all the 139 cultivated *I. batatas* accessions for the IGS marker. We thus used cpDNA lineage information from Roullier *et al.* [29] to complete our dataset. As described in Roullier *et al.* [29], nuclear markers show a more clear-cut geographic structure than do chloroplast markers. While the nuclear-marker-based cluster K1 is mostly composed of sweet potato accessions with Southern chloroplast haplotypes (33 out of 44), the nuclear-marker-based cluster K4 is mostly composed of sweet potato accessions bearing the Northern chloroplast haplotypes (14 out of 16), and the nuclear-marker-based cluster K3 includes sweet potatoes with both Northern and Southern chloroplast haplotypes (16 and 19 respectively). *Ipomoea* sp. specimens that group with the *I. trifida* cluster K2 carry the Northern chloroplast haplotype (or the unclassified rare haplotype 1).

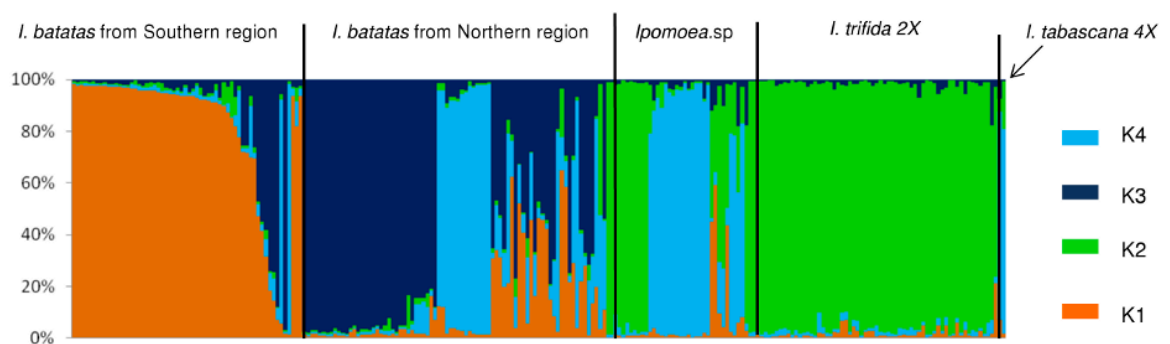
### *Genetic diversity of I. batatas, I. trifida, I. triloba and Ipomoea sp.*

The number of alleles  $NA$ , rarefied allelic richness  $Ar$ , and expected heterozygosity  $He$  were similar in *I. trifida*, *I. batatas* and *Ipomoea* sp. groups (Table 4-2), which also shared the same alleles. *Ipomoea triloba* showed fewer alleles than the other species (on average 3.8 alleles per locus, as compared to 9.5 to 12.6 alleles per locus for *I. trifida*, *I. batatas* and *Ipomoea* sp.). All diversity indices calculated ( $Ar$ ,  $He$  and  $Ho$ ) showed the same trend. The low values for both  $Ho$  and intra-taxon differentiation ( $D$ ) in *I. triloba* suggest the presence of null alleles in this species, all the more so since *I. triloba* alleles were quite different from those amplified for the other species. Although the most frequent alleles were shared among all groups, four “specific” alleles were present at a frequency greater than 0.1 in *I. triloba* and less than 0.02 in the other groups. Intra-taxon differentiation (mean pairwise Lynch distance between genotypes) was lower in *I. batatas* (0.424) than in *Ipomoea* sp. (0.566) and in *I. trifida* (0.716).

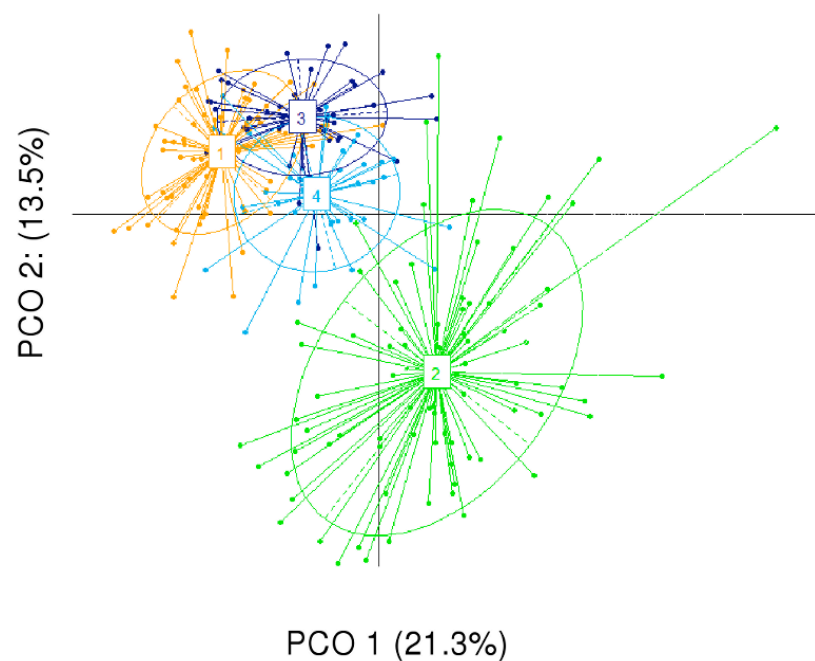
**Figure 4-4: Taxa boundaries of *I. batatas*, *I. trifida* and polyploid *Ipomoea* sp. accessions assessed using two complementary clustering approaches (SSRs dataset).**

a) Proportion of ancestry shared within each cluster, as estimated by the model-based clustering method implemented in STRUCTURE software. Each individual is represented as a vertical bar, with colours corresponding to probabilities of assignment to the four clusters. b) Principal Coordinate Analysis based on the Lynch distance between genotypes. Colours correspond to the clusters, as determined by the non model-based K-means hierarchical clustering method.

a)



b)



## DISCUSSION

### *The botanical origin of Ipomoea batatas*

Two scenarios were previously proposed for the formation of the *I. batatas* polyploid genome: autopolyploidization from *I. trifida* [26], or allopolyploidization involving two distant species [25]. The autopolyploidization scenario assumes *I. trifida* to represent an autopolyploid complex, with different ploidy levels (from diploid to hexaploid) from which cultivated *I. batatas* derived. However, cytological and marker-based studies suggested that the *I. batatas* hexaploid genome may be composed of two closely related genomes and a third one from a more distant relative [19,48]. Austin [25] postulated that the wild ancestor of *I. batatas* was a hybrid between *I. trifida* and *I. triloba* (allopolyploid scenario).

Chloroplast sequences and nuclear SSRs identified *I. trifida* as being more closely related to sweet potato than is *I. triloba* (Figures 4-2 and 4-3). On the basis of monoparentally inherited chloroplast sequences, *I. trifida* and *I. batatas* form a strongly supported monophyletic clade, demonstrating that at least one of the genomes of *I. batatas* is closely related to *I. trifida*. For *I. triloba*, both the low number of SSR alleles encountered and the strong deficit of heterozygotes suggest that null alleles were frequent in this species (Table 4-2). The low transferability of SSR markers between *I. batatas* and *I. triloba*, together with their strong divergence for nuclear and chloroplast sequences, show that *I. triloba* is more distantly related to sweet potato than is *I. trifida*. The overall similarity in allelic composition of *I. trifida* and *I. batatas* for nuclear SSRs, the low resolution of the ITS tree and the lower intra-taxon differentiation in *I. batatas* than in *I. trifida* (as measured by the mean pairwise distance between genotypes), all support an autopolyploid origin of *I. batatas*.

Chloroplast sequences (Figure 4-2), nuclear sequences (Figure 4-3a) and nuclear SSRs (Figures 4-3b, 4-3c and 4-4) all confirm the clear distinction between the two sweet potato gene pools previously identified on the basis of chloroplast and nuclear SSRs [29]. While both gene pools share an ancestry with *I. trifida*, their chloroplast sequences show more divergence from each other than from *I. trifida*, testifying to chloroplast ancestral polymorphism in the common ancestor of both species. These results point towards at least a double origin of *I. batatas*, with both gene pools sharing one ancestor with *I. trifida*.

Autopolyploidy has traditionally been considered to be the duplication of identical genomes. It is now appreciated that multiple origin of autopolyploidy is common [7,46,49,50]. Autopolyploid complexes may evolve by multiple independent genome duplication events [49], or result from the hybridization and genome doubling of differentiated populations of the same species, brought into contact, for example, by climate-induced range shifts [50]. Both scenarios should be considered for the origin of the *I. batatas* autopolyploid complex (Text S2).

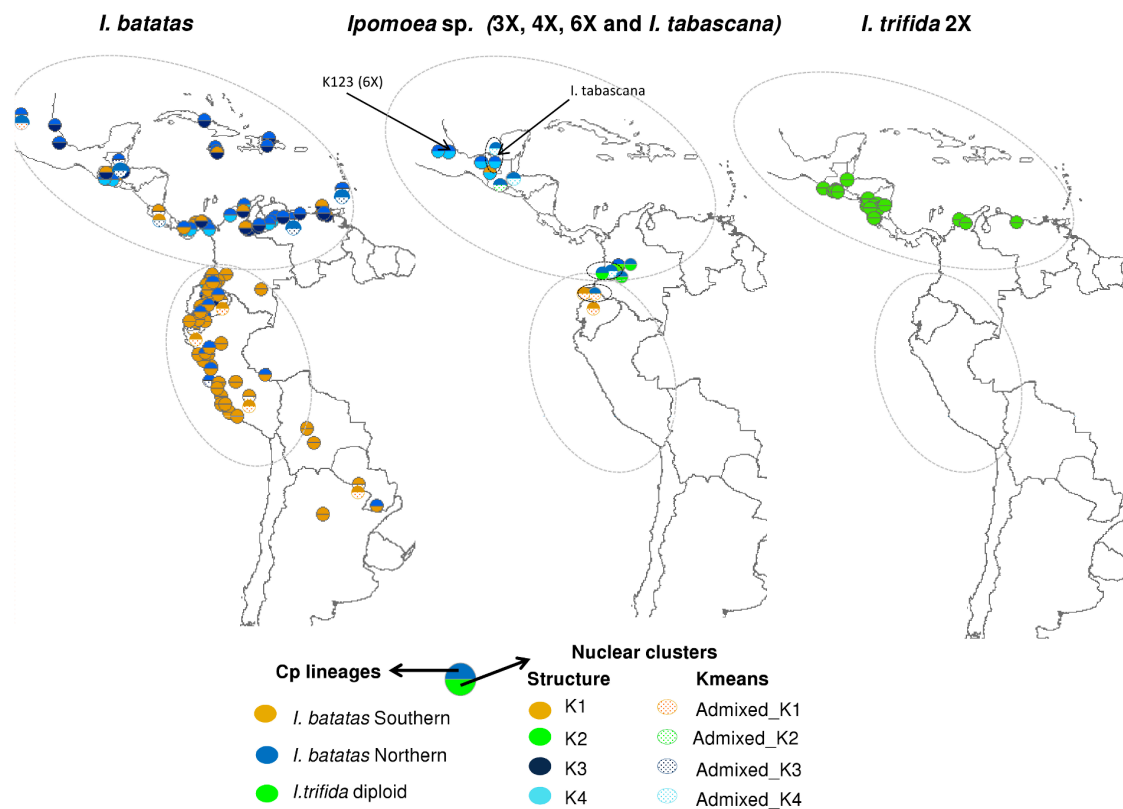
The absence of shared haplotypes for nuclear and chloroplast sequences between *I. trifida* and any *I. batatas* gene pool further suggests that diploid *I. trifida* thus cannot be considered as the direct wild progenitor of cultivated *I. batatas*. The closest wild relatives of domesticated sweet potato may instead be found among the polyploid accessions of dubious taxonomic status.

### *Samples of dubious taxonomical status and the origin of sweet potato*

At least two polyploidization/hybridization events, implying the existence of triploid or tetraploid intermediates, are necessary to explain how hexaploid sweet potato could have been generated from diploid progenitors (Text S2). Kobayashi [26] proposed that these intermediates might resemble the wild polyploids that have been collected from Mexico to northern Peru, accessions first identified as *I. trifida*. Taxonomical revision subsequently placed most of them in *I. batatas* [27]. This latter re-identification was supported by the fact that these wild *Ipomoea* sp. accessions were more easily crossed with *I. batatas* than with diploid *I. trifida* [51]. Our findings, based on both nuclear and chloroplast markers, show that diploid *I. trifida* indeed form a monophyletic group, exclusive of these *Ipomoea* sp. accessions of higher ploidy.

**Figure 4-5: Geographical patterns of cpDNA lineages and nuclear clusters (STRUCTURE results) of *I. batatas*, *I. trifida* and polyploid *Ipomoea* sp. accessions.**

The top half of the circle provides the chloroplast lineage while the bottom half gives the nuclear genome as revealed by STRUCTURE grouping. For admixed samples, K-means clusters are provided. Each circle represents one accession, unless samples of the same accession provided different information. In the latter case, samples of the same accession are surrounded.



Among the polyploid *Ipomoea* sp. accessions used in our study, numerous ones shared haplotypes with *I. batatas*, but none with *I. trifida*. The few *Ipomoea* sp. accessions collected in Ecuador carried a Southern *I. batatas* chloroplast haplotype, while all other *Ipomoea* sp. accessions carried the Northern *I. batatas* chloroplast haplotype. Nuclear SSRs placed *Ipomoea* sp. accessions as intermediate between diploid *I. trifida* and cultivated *I. batatas*. While four accessions (seven samples) of *Ipomoea* sp., all from Colombia, were clearly assigned to the diploid *I. trifida* cluster K2, all others grouped with *I. batatas* Northern cluster K4 (16 samples from Mexico and Guatemala) or were considered as “admixed” between *I. trifida* and *I. batatas* K1 (three samples from Ecuador) or K4 (two samples from Mexico, two from Guatemala, one from Colombia) nuclear clusters. The seven specimens indicated as *I. batatas* by both kinds of markers may represent original wild *I. batatas*, i.e., forms intermediate between the diploid progenitors and cultivated hexaploid *I. batatas* (as was strongly suggested by Austin [25]; and Bohac *et al.* [27]). The 15 specimens with admixed nuclear genomes, or with *I. trifida* nuclear genomes and *I. batatas* chloroplast haplotypes, may be feral individuals issued from crosses between cultivated hexaploid *I. batatas* (as the maternal parent) and diploid *I. trifida*. Like most *Ipomoea* sp. accessions, *I. tabascanana*, the putative hybrid between *I. batatas* and *I. trifida* [30], carries a nuclear genome admixed between *I. trifida* and *I. batatas* K4, consistent with this hybrid status.

Our genetic results confirmed previous taxonomic studies, which identify polyploid (3X, 4X and few 6X) *Ipomoea* sp. in the *I. batatas* species. They are additional evidence that *I. batatas* may exist not only as hexaploid cultigens, but also as a true wild species, with several ploidy levels (from 3X to 6X at least), forming in fact an autopolyploid complex with multiple origins. Recently, an analysis of Waxy intron variations argued for an allopolyploid origin for the sweet potato, which probably occurred by hybridizations between *I. tenuissima* and *I. littoralis* [52]. *I. littoralis* was described by Austin [53] to be the only species of the genus native and endemic to the Old World. Some wild tropical American tetraploid plants may have been misidentified as *I. littoralis* (Kobayashi, 1984). *I. littoralis* accessions used in the study by Gao *et al.* [52] likely indeed are specimens of wild *I. batatas*. Their results would then confirm ours in supporting the fact that *I. batatas* has multiple origins.

#### *How, when and where was sweet potato domesticated?*

It is quite clear that crop origins in Neotropical America were spatially diffuse, and often occurred in a number of localities in both tropical Central and South America [15]. Previous molecular analyses suggested that sweet potato may be Central American in origin. Results of a recent study [29] suggest that at least two domestications occurred, one in Caribbean/Central America, and one in northwestern South America, giving rise to two domesticated gene pools (the Northern and Southern ones). Our study confirms this phylogeographic pattern, but does not allow us to pinpoint where domestication took place, because we still lack a representative sample of wild *I. batatas* populations. In one scenario, this autopolyploid complex may have evolved in distinct geographical areas (perhaps at the periphery of ranges of their diploid ancestors), raising the possibility that two independent groups of cultivators took advantage of distinct local wild polyploid *I. batatas* populations. Alternatively, the phylogeographic pattern found within *I. batatas* might result from post-domestication patterns of human expansion. In this second scenario, a single domestication from polymorphic wild polyploid populations (resulting from independent, geographically restricted autopolyploidization events or from the contact, hybridization and genome doubling of pre-differentiated populations) may have occurred (Figure 4-6). Considering all available evidence, the first scenario seems to us the more parsimonious (Supplementary material 3).

Sweet potato is one of the oldest domesticates in the Americas, with archaeological remains of dried sweet potatoes from Peru dating back to 10,000 to 8,000 BP [15,54]. Unlike other plants (maize, wheat or pearl millet, for example), cultivated *I. batatas* has not suffered a severe bottleneck during domestication: the crop’s nuclear genetic diversity, at least as estimated based on neutral markers, is comparable to that of its polyploid progenitor, and only slightly lower than that of diploid *I. trifida* in terms of number of alleles. Autopolyploid formation allows a large part of the genetic diversity present in progenitors to be incorporated, particularly when multiple maternal lineages are added [7]. However, it is also possible that the domestication process in itself resulted in retention of a large part of the

**Table 4-1: Contingency table confronting cpDNA haplotype lineages with Structure clusters among *I. batatas*, *I. trifida*, *Ipomoea* sp. and *I. tabascana*.**

Ancestry for nuclear SSRs, as inferred by STRUCTURE	cp DNA haplotype lineages				
	Southern <i>I. batatas</i>		Northern <i>I. batatas</i>		Diploid <i>I. trifida</i>
	<i>I. batatas</i>	<i>Ipomoea</i> . sp	<i>I. batatas</i>	<i>Ipomoea</i> . sp	<i>I. tabascana</i>
K1 (Southern <i>I. batatas</i> )	33		11		
K2 ( <i>I. trifida</i> )				7	65
K3 (Northern <i>I. batatas</i> )	19		16		
K4 (Northern <i>I. batatas</i> )	1	2	14	5	1
Admixed “ <i>I. trifida</i> / <i>I. batatas</i> ”		3	1	5	
Admixed “ <i>I. batatas</i> / <i>I. batatas</i> ”	17		21		1

“Admixed *I. trifida* / *I. batatas*” correspond to individuals sharing ancestry (< 80 %) for nuclear SSRs with K2 and any of the clusters K1, K3 or K4.

“Admixed *I. batatas* / *I. batatas*” share ancestry for nuclear SSRs between any two of the three *I. batatas* clusters K1, K3 or K4.

diversity of ancestral wild populations. The major trait of sweet potato's domestication syndrome is the development of edible tuberous storage roots, a trait with complex determinism [55], likely linked to polyploidy. Indeed, diploid *I. trifida* do not produce edible roots, but often present some thickened ones. Wild tetraploid *I. batatas* form only thickened "pencil-shaped" roots [27] and wild hexaploid *I. batatas* populations have been too scantily characterized to provide any reliable description of their roots and tubers. Even in the cultivated forms, tuberization is a labile trait, with complex genetic and environmental determinism and low heritability [51,55]. Moreover, gene flow between the crop and wild relatives (*I. trifida*, and more particularly wild *I. batatas*) is still possible in natural settings despite their different ploidy levels, but artificial crosses show reduced seed set and poor yields [51,56]. With current data, it is not possible to establish a tight framing of the timing of domestication relative to polyploidization. We postulate, however, that during the domestication process, cultivators may have repeatedly captured and multiplied tuber-bearing wild mutants or crop/wild hybrids, probably hexaploid plants. Crop/wild gene flow would have progressively decreased under cultivation, favouring recombination between tuber-bearing cultivated forms and stabilizing the formation and development of storage roots.



**Table 4-2: Genetic diversity of the four geographically well-sampled taxa as revealed by nuclear SSRs.**

Taxa	<i>I. batatas</i> (139)				<i>I. trifida</i> (65)				<i>Ipomoea</i> sp. (39)				<i>I. triloba</i> (26)		
Locus	NA	Ar	He	Ho	NA	Ar	He	Ho	NA	Ar	He	Ho	NA	He	Ho
<b>J206A</b>	9	6.5	0.70	0.95	9	7.5	0.81	0.6	7	6.8	0.81	0.83	3	0.58	0.11
<b>J544b</b>	8	7.1	0.74	0.97	12	10.4	0.84	0.52	7	6.8	0.72	0.75	5	0.73	0.19
<b>J263</b>	7	5	0.73	0.91	7	5.9	0.73	0.35	4	4	0.59	0.70	2	0.18	0.11
<b>J116A</b>	15	13	0.83	0.99	10	8.5	0.81	0.61	9	8.8	0.82	0.81	3	0.56	0.04
<b>IbS11</b>	13	10	0.83	0.99	15	11.1	0.87	0.66	9	8.8	0.83	0.97	1	0.42	0.16
<b>IbC5</b>	13	10	0.80	0.97	12	9.6	0.87	0.66	10	9.9	0.84	0.89	8	0.86	0.12
<b>ib297</b>	24	14.9	0.85	0.74	30	20.3	0.97	0.44	18	17.1	0.9	0.64	6	0.79	0.08
<b>J522A</b>	10	6.4	0.74	0.94	6	5.4	0.80	0.54	12	11.2	0.76	0.86	2	0.46	0.04
<i>Mean</i>	<i>12.4</i>	<i>9.1</i>	<i>0.78</i>	<i>0.93</i>	<i>12.62</i>	<i>9.8</i>	<i>0.84</i>	<i>0.55</i>	<i>9.5</i>	<i>9.2</i>	<i>0.78</i>	<i>0.81</i>	<i>3.75</i>	<i>0.57</i>	<i>0.1</i>
<i>D</i>	0.424				0.716				0.566				0.36		

Values for the number of alleles (*NA*), its rarefied value over 25 individuals (1000 resamplings; *Ar*) and the observed and expected heterozygosities *Ho* and *He*, are provided both per locus and as mean values averaged over all loci. *D* corresponds to the intra-taxon mean Lynch distance between genotypes.

## CONCLUSION

Previous cytogenetic and neutral-marker-based studies have pointed to diploid *I. trifida* as the closest wild relative of sweet potato. We argue here that polyploid true wild *I. batatas* populations of multiple autopolyploid origins exist. Their diploid progenitors have yet to be identified. Sharing ancestry with extant diploid *I. trifida*, these putative diploid ancestors might be extinct or may simply have not been collected. Wild tuber-bearing populations (most probably hexaploids) may have been domesticated independently in South America and the Caribbean/Central America. These two gene pools have secondarily come into contact, as shown by their admixture for both chloroplast and nuclear markers.

To obtain further insights into where, how, and when polyploidization and sweet potato domestication occurred, marker-based studies should consider both homologous (similar gene copies within a progenitor diploid genome) and homeologous (similar gene copies brought together in a polyploid genome after the hybridization and genome doubling of differentiated diploid progenitors) copies of domestication genes, with a comprehensive sampling of both wild and cultivated *I. batatas*.

Finally, a simplified vision of domestication, which considers wild and domesticate as complementary rather than antonymous [57], may have conduced geneticists, working on the origin of sweet potato, to miss the importance of wild intermediate forms of *I. batatas*. However, taxonomists have described them for a long time [27,58]. Further elucidation of the cultivation history of the crop will require a better understanding of the distribution and ecology of these wild, cultivated and intermediate forms of *I. batatas* throughout the Neotropics, and their consideration in future genetic studies.

## ACKNOWLEDGEMENTS

The authors are very grateful to Austin D. F. (from the Arizona-Sonora Desert Museum) for useful comments on previous versions of the manuscript and for stimulating discussions. We also thank CIP field technicians for providing plant material. Data used in this work were produced through the technical facilities of the Centre Méditerranéen Environnement Biodiversité.

## REFERENCES

- Otto SP (2007) The evolutionary consequences of polyploidy. *Cell* 131: 452-462.
- Purugganan MD, Fuller DQ (2009) The nature of selection during plant domestication. *Nature* 457: 843-848
- Zohary D (2004) Unconscious selection and the evolution of domesticated plants. *Econ Bot* 58: 5-10.
- Allaby RG, Fuller DQ, Brown TA (2008) The genetic expectations of a protracted model for the origins of domesticated crops. *Proc Natl Acad Sci* 105: 13982-13986.
- McKey D, Elias M, Pujol B, Duputié A (2010) The evolutionary ecology of clonally propagated domesticated plants. *New Phytol* 186: 318-332.
- Emshwiller E (2006) Origins of polyploid crops - The example of the octoploid tuber crop *Oxalis tuberosa*. In: Zeder MA, Bradley DG, Emshwiller E, Smith BD, editors. Documenting domestication: New genetic and archaeological paradigms. Berkeley, CA: University of California Press. pp. 153-168.
- Parisod C, Holderegger R, Brochmann C (2010) Evolutionary consequences of autopolyploidy. *New Phytol* 186: 5-17.
- Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, et al. (2009) The frequency of polyploid speciation in vascular plants. *Proc Natl Acad Sci* 106: 13875-13879.
- Robertson A, Rich TCG, Allen AM, Houston L, Roberts C, et al. (2010) Hybridization and polyploidy as drivers of continuing evolution and speciation in *Sorbus*. *Mol Ecol* 19: 1675-1690.
- Hilu KW (1993) Polyploidy and the evolution of domesticated plants. *Am J Bot* 80: 1494-1499.
- Perrier X, De Langhe E, Donohue M, Lentfer C, Vrydaghs L, et al. (2011) Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proc Natl Acad Sci* 108: 11311-11318.
- Pickersgill B (2007) Domestication of plants in the Americas: insights from Mendelian and molecular genetics. *Ann Bot* 100: 925-940.
- Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez G J, Buckler E, et al. (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proc Natl Acad Sci* 99: 6080-6084.
- Ranere AJ, Piperno DR, Holst I, Dickau R, Iriarte J (2009) The cultural and chronological context of early Holocene maize and squash domestication in the Central Balsas River Valley, Mexico. *Proc Natl Acad Sci* 106: 5014-5018.
- Piperno DR (2011) The origins of plant cultivation and domestication in the New World tropics. *Curr Anthropol* 52: 453-470.
- Chacón MI, Pickersgill B, Debouck DG (2004) Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. *Theor Appl Genet* 110: 432-444.
- Motta-Aldana JR, Serrano-Serrano ML, Hernández-Torres J, Castillo-Villamizar G, Debouck DG, et al. (2010) Multiple origins of Lima bean landraces in the Americas: Evidence from chloroplast and nuclear DNA polymorphisms. *Crop Sci* 50: 1773-1787.
- Lebot V (2009) Tropical Root and Tuber Crops - Cassava, Sweet Potato, Yams and Aroids. Wallingford, UK: CABI Publishing Group.
- Buteler MI, Jarret RL, LaBonte DR (1999) Sequence characterization of microsatellites in diploid and polyploid *Ipomoea*. *Theor Appl Genet* 99: 123-132.
- Huang J, Corke H, Sun M (2002) Highly polymorphic AFLP markers as a complementary tool to ITS sequences in assessing genetic diversity and phylogenetic relationships of sweetpotato (*Ipomoea batatas* (L.) Lam.) and its wild relatives. *Genet Resour Crop Evol* 49: 541-550.
- Huang J, Sun M (2000) Genetic diversity and relationships of sweetpotato and its wild relatives in *Ipomoea* series *Batatas* (Convolvulaceae) as revealed by inter-simple sequence repeat (ISSR) and restriction analysis of chloroplast DNA. *Theor Appl Genet* 100: 1050-1060.
- Jarret RL, Austin DF (1994) Genetic diversity and systematic relationships in sweetpotato (*Ipomoea batatas* (L.) Lam.) and related species as revealed by RAPD analysis. *Genet Resour Crop Evol* 41: 165-173.
- Rajapakse S, Nilmalgoda SD, Molnar M, Ballard RE, Austin DF, et al. (2004) Phylogenetic relationships of the sweetpotato in *Ipomoea* series *Batatas* (Convolvulaceae) based on nuclear beta-amylase gene sequences. *Mol Phylogenet Evol* 30: 623-632.
- Srisuwan S, Sihachakr D, Siljak-Yakovlev S (2006) The origin and evolution of sweet potato (*Ipomoea batatas* Lam.) and its wild relatives through the cytogenetic approaches. *Plant Sci* 171: 424-433.
- Austin DF. The taxonomy, evolution and genetic diversity of sweet potatoes and related wild species; 1988; Lima, Peru. International Potato Center (CIP). pp. 27-58.
- Kobayashi M (1984) The *Ipomoea trifida* complex closely related to sweet potato. In: Shideler SF, Rincon H, editors. Proceedings of the 6th Symposium of the International Society of Tropical Root Crop. Lima, Peru: CIP. pp. 561-568.
- Bohac JR, Austin DF, Jones A (1993) Discovery of wild tetraploid sweetpotatoes. *Econ Bot* 47: 193-201.

28. Wendel JF, Schnabel A, Seelanan T (1995) Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proc Natl Acad Sci* 92: 280.
29. Roullier C, Rossel G, Tay D, McKey D, Lebot V (2011) Combining chloroplast and nuclear microsatellites to investigate origin and dispersal of New World sweet potato landraces. *Mol Ecol* 20: 3963-3977.
30. McDonald JA, Austin DF (1990) Changes and additions in *Ipomoea* section *Batatas* (Convolvulaceae). *Brittonia* 42: 116-120.
31. Komaki K, Regmi H, Katayama K, Tamiya S (1998) Morphological and RAPD pattern variations in sweetpotato and its closely related species. *Breed Sci* 48: 281-286.
32. Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Am J Bot* 94: 275-288.
33. White TJ, Bruns T, Lee SB, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, editors. *PCR Protocols*. San Diego: Academic Press. pp. 315-322.
34. Ingvarsson PK, Ribstein S, Taylor DR (2003) Molecular evolution of insertions and deletion in the chloroplast genome of *Silene*. *Mol Biol Evol* 20: 1737-1740.
35. Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Syst Biol* 49: 369-381.
36. Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754-755.
37. Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9: 1657-1659.
38. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23: 254-267.
39. Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289-290.
40. Perrier X, Jacquemoud-Collet JP (2006) DARwin software pp. <http://darwin.cirad.fr/>
41. Clark LV, Jasieniuk M (2011) POLYSAT: an R package for polyploid microsatellite analysis. *Mol Ecol Resour* 11: 562-566.
42. Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes* 7: 574-578.
43. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
44. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14: 2611-2620.
45. Arrigo N, Felber F, Parisod C, Buerki S, Alvarez N, et al. (2010) Origin and expansion of the allotetraploid *Aegilops geniculata*, a wild relative of wheat. *New Phytol* 187, 1170-1180.
46. Burnier J, Buerki S, Arrigo N, Küpfer P, Alvarez N (2009) Genetic structure and evolution of Alpine polyploid complexes: *Ranunculus kuepferi* (Ranunculaceae) as a case study. *Mol Ecol* 18: 3730-3744.
47. Arrigo N, Guadagnuolo R, Lappe S, Pasche S, Parisod C, et al. (2011) Gene flow between wheat and wild relatives: empirical evidence from *Aegilops geniculata*, *Ae. neglecta* and *Ae. triuncialis*. *Evol Appl* 4: 685-695.
48. Magoon ML, Krishnan R, Vijaya Bai K (1970) Cytological evidence on the origin of sweet potato. *Theor Appl Genet* 40: 360-366.
49. Bardy KE, Albach DC, Schneeweiss GM, Fischer Ma, Schönswetter P (2010) Disentangling phylogeography, polyploid evolution and taxonomy of a woodland herb (*Veronica chamaedrys* group, Plantaginaceae s.l.) in southeastern Europe. *Mol Phylogenet Evol* 57: 771-786.
50. Parisod C, Besnard G (2007) Glacial in situ survival in the Western Alps and polytopic autopolyploidy in *Biscutella laevigata* L.(Brassicaceae). *Mol Ecol* 16: 2755-2767
51. Shiotani I, Yoshida S, Kawase T (1990) Numerical taxonomic analysis and crossability of diploid *Ipomoea* species related to the sweet potato. *Japanese Journal of Breeding* 40: 159-174.
52. Gao M, Ashu GM, Stewart L, Akwe Wa, Njiti V, et al. (2012) Wx intron variations support an allohexaploid origin of the sweetpotato [*Ipomoea batatas* (L.) Lam.]. *Euphytica* 177: 111-133.
53. Austin DF (1991) *Ipomoea littoralis* (Convolvulaceae)-Taxonomy, Distribution, and Ethnobotany. *Econ Bot* 45: 251-256.
54. Ugent D, Peterson W (1988) Archeological remains of potato and sweet potato in Peru. *CIP Circular* 16: 1-10.
55. Ravi V, Naskar S, Makeshkumar T, Babu B, Krishnan BSP (2009) Molecular physiology of storage root formation and development in sweet potato (*Ipomoea batatas* (L.) Lam.). *Journal of Root Crops* 35: 1-27.

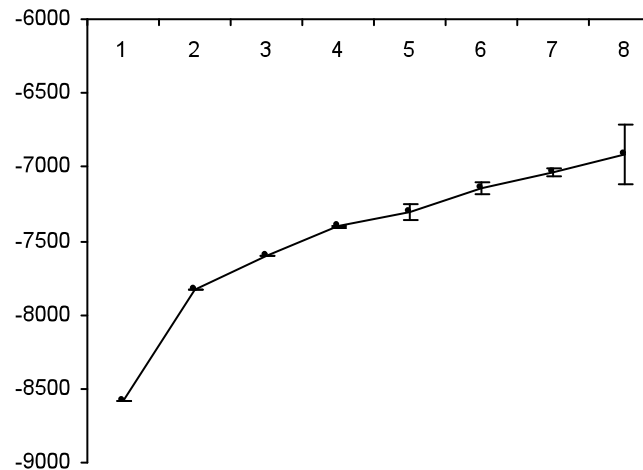
56. Orjeda G, Freyre R, Iwanaga M (1991) Use of *Ipomoea trifida* germ plasm for sweet potato improvement. 3. Development of 4x interspecific hybrids between *Ipomoea batatas* (L.) Lam ( $2n=6x=90$ ) and *I. trifida* (H.B.K) G. Don ( $2n=2x=30$ ) as storage-root initiators for wild species. Theor Appl Genet 83: 159-163.
57. Dobney K, Larson G (2006) Genetics and animal domestication: new windows on an elusive process. J Zool 269: 261-271.
58. Jones A (1967) Should Nishiyama's K123 (*Ipomoea trifida*) be designated *I. batatas*? Econ Bot 21: 163-166.

**Supplementary material 1: Screening for the optimal number of clusters for Structure (a and b) and K-means clustering results (c and d).**

For Structure results:

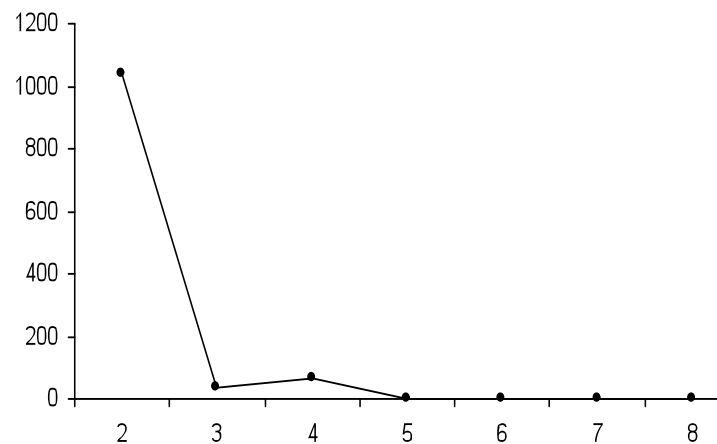
**a) Variation of the posterior log-probability of the data as a function of the number of clusters  $K$ .**

Values of likelihood increased from  $K = 1$  to  $K = 8$ , showing that the fit of the model to the data is continuously improved when the number of clusters is increased.



**b) Variation of  $\Delta K$  values [1].**

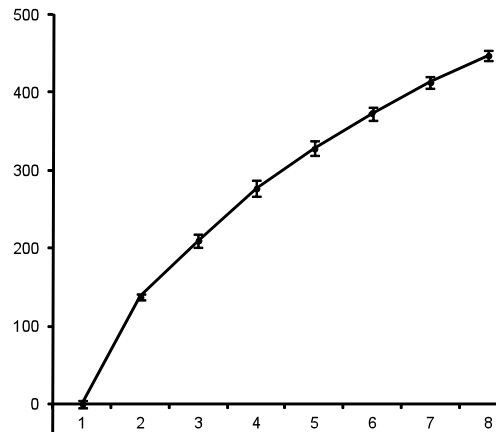
The appropriate number of clusters was unclear. We finally retained a  $K$  value of 4, value that provided the least admixed genotypes and that we considered the most relevant to describe our dataset.



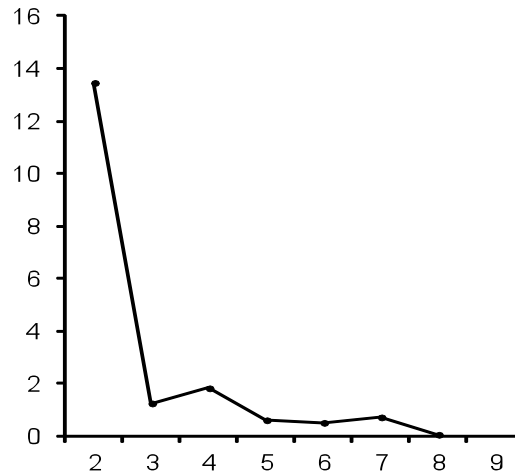
### **For K-means data:**

For screening the “optimal” number of groups  $K$ , we used an adaptation of the  $\Delta K$  criterion [1] proposed by Arrigo *et al.* [2], with intergroup inertia as a proxy of clustering accuracy.

### **c) Variation of the inertia**



**d) Variation of  $\Delta K$  values showed one major peak at respectively  $K = 2$  and a minor peak at  $K = 4$ .**



### **References cited:**

1. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14: 2611-2620.
2. Arrigo N, Felber F, Parisod C, Buerki S, Alvarez N, et al. (2010) Origin and expansion of the allotetraploid *Aegilops geniculata*, a wild relative of wheat. *New phytol* 187, 1170-1180.



## Supplementary material 2: Fuzzy c-means clustering method and results.

Structure is a model-based algorithm, assuming Hardy-Weinberg equilibrium, to assign individuals within a pre-defined number of genetic groups [1]. As it is unclear whether model-based clustering methods are relevant to address data in the context of polyploidy, especially when microsatellite alleles must be treated as dominant markers (binary data), we confronted this analysis with two non model-based clustering approaches, hierarchical K-means clustering and fuzzy c-means clustering methods.

We first performed a non-hierarchical Kmeans clustering. This technique assigns individuals to  $K$  genetic groups, so as to maximize intergroup variance (quantified by inertia). We also performed a fuzzy c-means clustering analysis [2]. This algorithm allocates individuals to  $K$  groups using an iterative process, as does Kmeans clustering. It also calculates probability of membership of an individual in each group. These membership probabilities can be “compared” to ancestry values inferred by Structure. Their computations rely on the fuzzification parameter  $r$  (ranging between one and infinity) determining the stringency of clustering: the larger the value of  $r$ , the fuzzier the clustering is. Because no specific recommendations exist for properly setting  $r$  values, we tested values for  $r$  ranging from 1.1 to 1.5, previous studies generally using  $r$  values ranging between 1 and 2 [2,3]. In our case, only low values of  $r$  gave relevant grouping results. For values superior to 1.25, all individuals were “admixed”, with a membership value to any cluster  $< 80\%$ . We show here results obtained for  $r = 1.1$ .

All different clustering methods globally gave congruent results, as attested by the contingency table (Figure 1). C-means provided “fuzzier” results than STRUCTURE: a larger number of individuals were not clearly « assigned » to a given cluster (probability of membership in any cluster  $< 80\%$ ) (Figure 2)

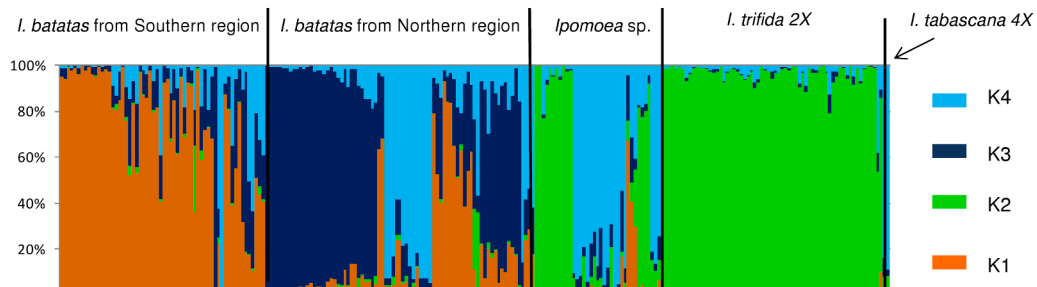
**Figure 1. Contingency table confronting the clustering results obtained through Kmeans and fuzzy c-means and Structure analyses at  $K = 4$ .**

Admixed corresponds to individuals with membership probabilities to any cluster  $< 80\%$ .

	Structure clusters				
	K1	K2	K3	K4	Admixed
Kmeans clusters					
K1	47		3		20
K2		74			8
K3			34		14
K4				34	9
c-means clusters					
K1	32				6
K2		73			
K3			30		4
K4				23	1
Admixed	15	1	7	11	40

**Figure 2. Probabilities of membership in each cluster, as estimated by the non model-based fuzzy c-means clustering method.**

Each individual is represented as a vertical bar, with colours corresponding to probabilities of membership in each cluster (K1, K2, K3 and K4).

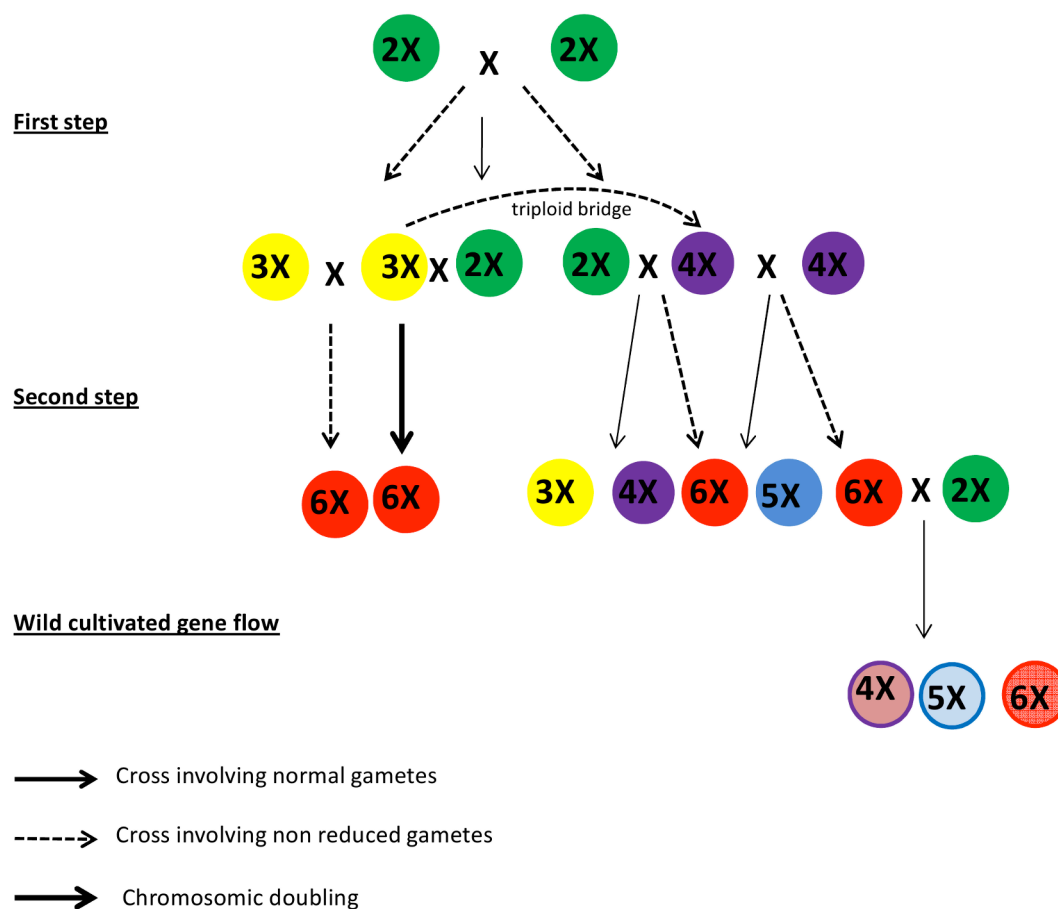


### References cited

1. Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7: 574-578.
2. Arrigo N, Guadagnuolo R, Lappe S, Pasche S, Parisod C, et al. (2011) Gene flow between wheat and wild relatives: empirical evidence from *Aegilops geniculata*, *Ae. neglecta* and *Ae. triuncialis*. Evol Appl 4: 685-695.
3. Gompert Z, Lucas LK, Fordyce JA, Forister ML, Nice CC (2010) Secondary contact between *Lycaeides idas* and *L. melissa* in the Rocky Mountains: extensive admixture and a patchy hybrid zone. Mol Ecol 19: 3171-3192.

### Supplementary material 3: Scenarios about the origins of *Ipomoea batatas* polyploidization and domestication

The formation of the hexaploid genome must have involved at least two steps, from diploidy to intermediate ploidy levels (triploid or tetraploid) and then hexaploidy. The most likely polyploidization route in sweet potato involves sexual mechanisms *via* the production of  $2n$  gametes, whose occurrence has been demonstrated in diploid and triploid *I. trifida*, as well as in tetraploid *I. batatas* [1,2,3]. Moreover, polyploid *Ipomoea* sp. (mostly 4X) analyzed in our study may have two distinct origins: i) original intermediate wild forms of *I. batatas* (solid circle in Figure 1), or ii) feral plants issued from crosses between hexaploid *I. batatas* and a diploid wild relative (same but transparent colored circles in Figure 1).



**Figure 1: Mechanisms of formation of the hexaploid genome of sweet potato**

*Triploid bridge: formation of triploids is often considered to mediate the autopolyploidization process. It is also possible that tetraploids, as well as hexaploids, were formed without the intermediate formation of triploids, simply via successive crosses involving  $2n$  gametes (right part of the graph).*

In this study, we were unable to provide a tight framing of the location and the timing of polyploidization and domestication of sweet potato. This failure is mainly due to the fact that we still lack a good description (taxonomy, distribution and ecology) and sampling of wild and intermediate forms of *I. batatas* throughout the neotropics. Several scenarios might be considered, taking into account that both natural and human-mediated factors have modified the range of *I. batatas* and the distribution of genetic diversity within the species.

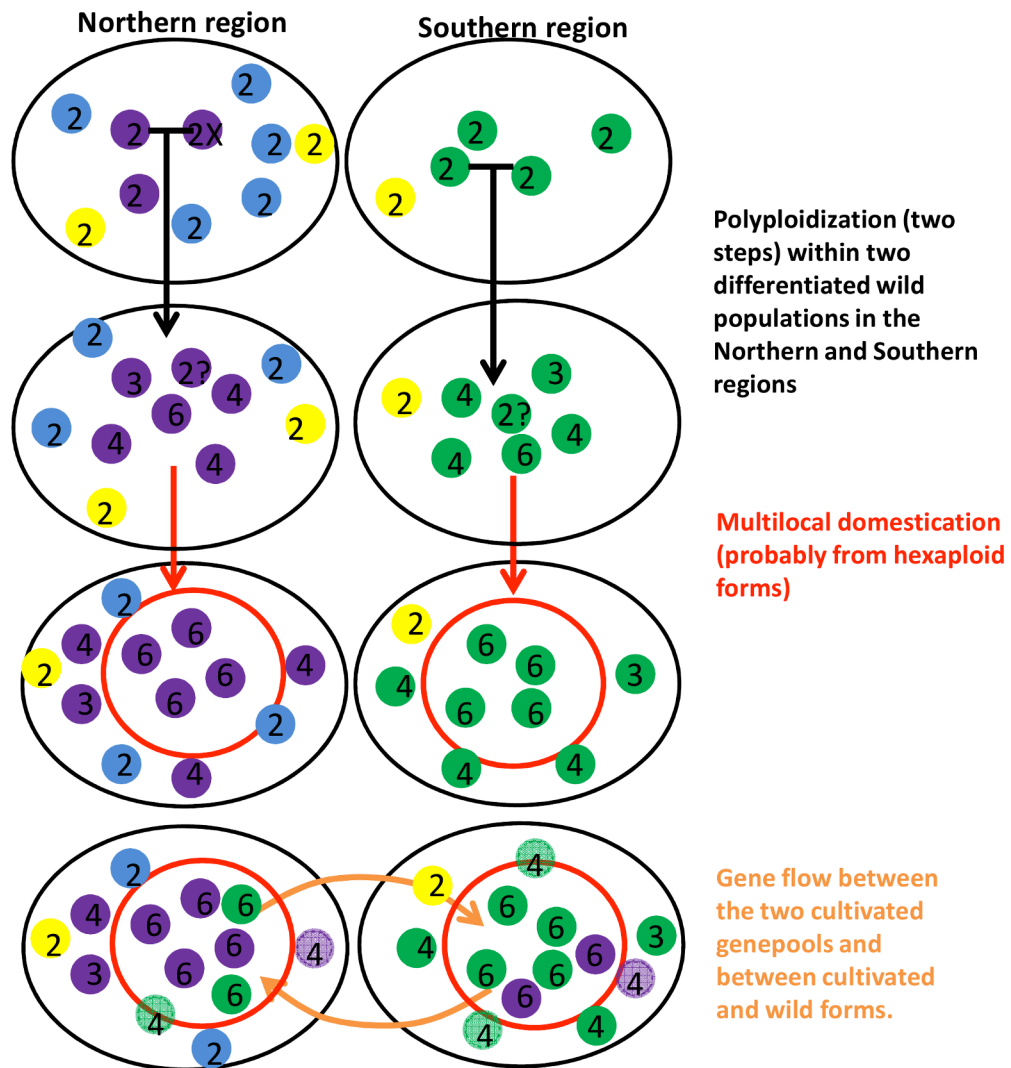
Autopolyploidy is more common in plants than is traditionally assumed. Our genetic results all support the hypothesis that *I. batatas* should be regarded as an autopolyploid complex with multiple origins. Extant diploid *I. trifida* (which extend from northwestern Mexico to Colombia and from sea level to 1200m) cannot be considered as the direct progenitor of *I. batatas*. Strict diploid progenitors have not yet been identified. Sharing ancestry with extant diploid *I. trifida*, these putative diploid ancestors might be extinct or have simply not yet been collected.

Intermediate tetraploid wild *I. batatas* have been collected through tropical America from southern Mexico to Ecuador (and from sea level to 1600m), and probably extend south at least to northern Peru. Wild hexaploid *I. batatas* have essentially been collected in Mexico. In this article we argue that domestication probably involved selection of hexaploids, whose high ploidy level likely “favours” tuberization. Most of the wild hexaploid progenitors might have been retained during the domestication process, through the successive capture and clonal propagation of tuber-bearing wild plants. Crop/wild gene flow would have progressively decreased under cultivation, favouring recombination between tuber-bearing cultivated forms and stabilizing the formation and development of storage roots.

We present here four different scenarios which may underlie genetic patterns described in the article. Scenario 1 is the most parsimonious considering available data. However, other scenarios are also presented to illustrate the diversity and complexity of possibilities.

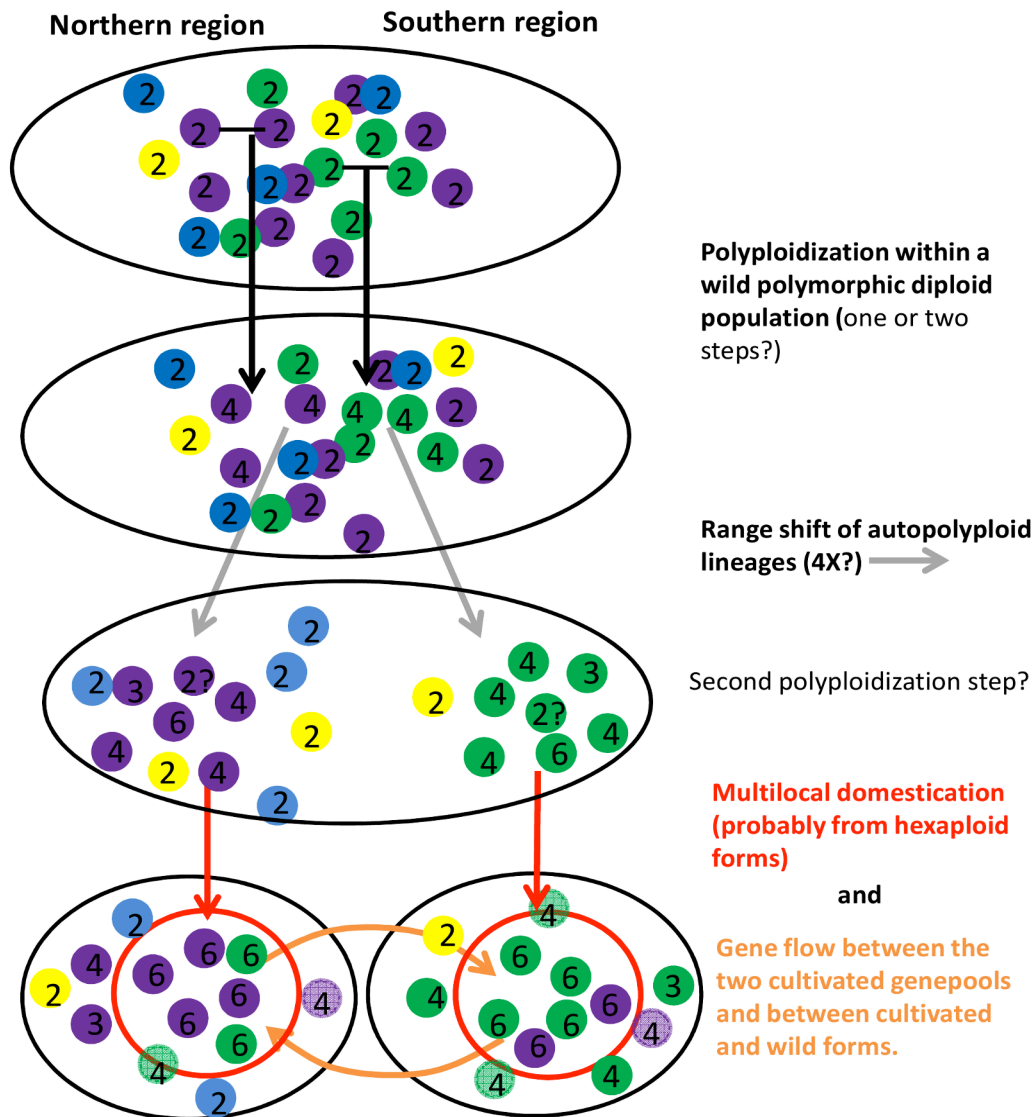
**Scenario 1:** The most parsimonious scenario, explaining the clear-cut phylogeographical pattern inferred from both nuclear and chloroplast data:

*Multiple independent events of autopolypoidy within several polymorphic and pre-differentiated wild populations (phylogeographical differentiation), and multi-local domestication within each polyploid population.*

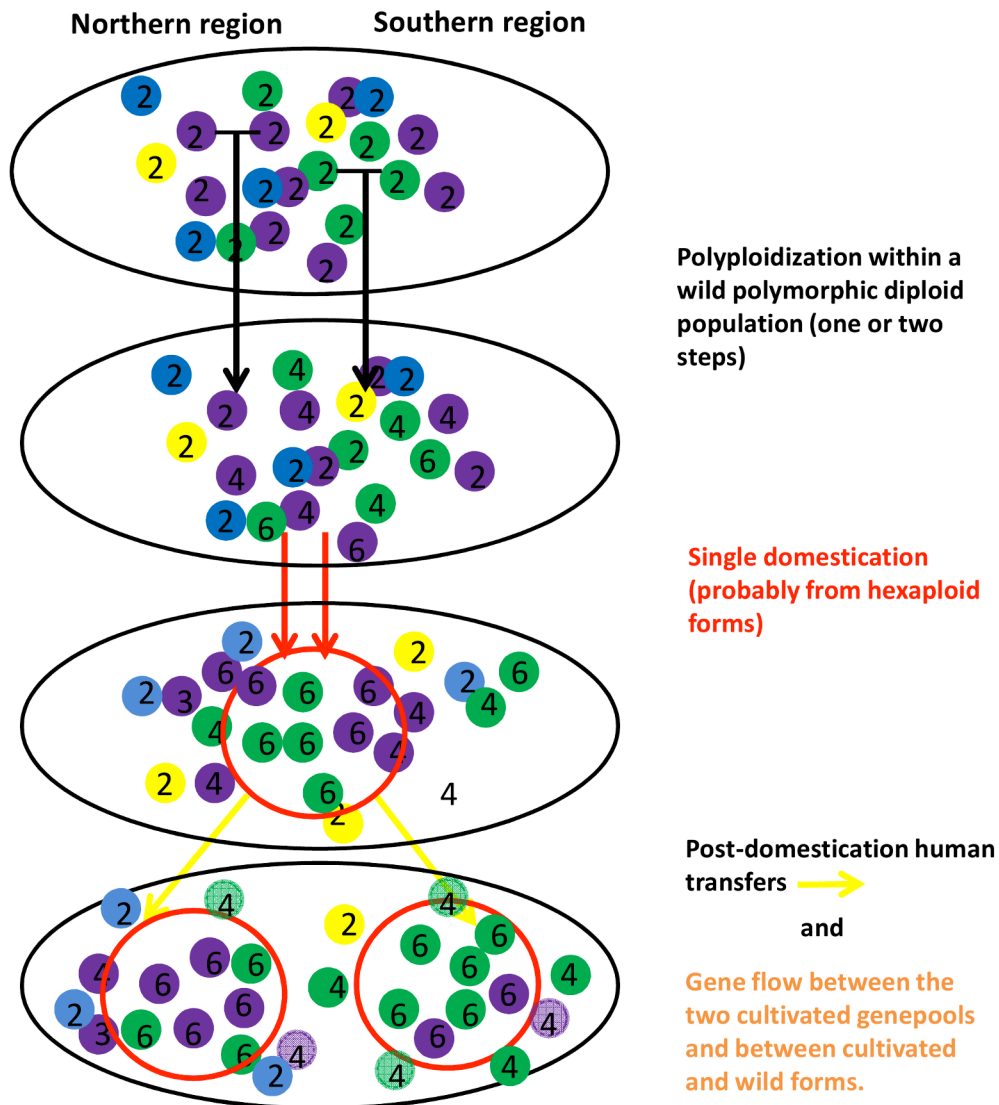


**Scenario 2:** Some biogeographical and ecological studies have highlighted an association between the formation of autopolyploid lineages and environmental changes [4]. In this context, autopolyploidy may be regarded as a mechanism enhancing colonization ability, and polyploids may spread out of the range of their progenitors. It is quite possible that:

*Multiple events of autopolyploidy occurred within a diploid polymorphic wild population, somewhere in the tropical American lowlands. The clear-cut phylogeographical pattern would then be inherited from the independent expansion of the range of both polyploid lineages and their subsequent multilocal domestications.*



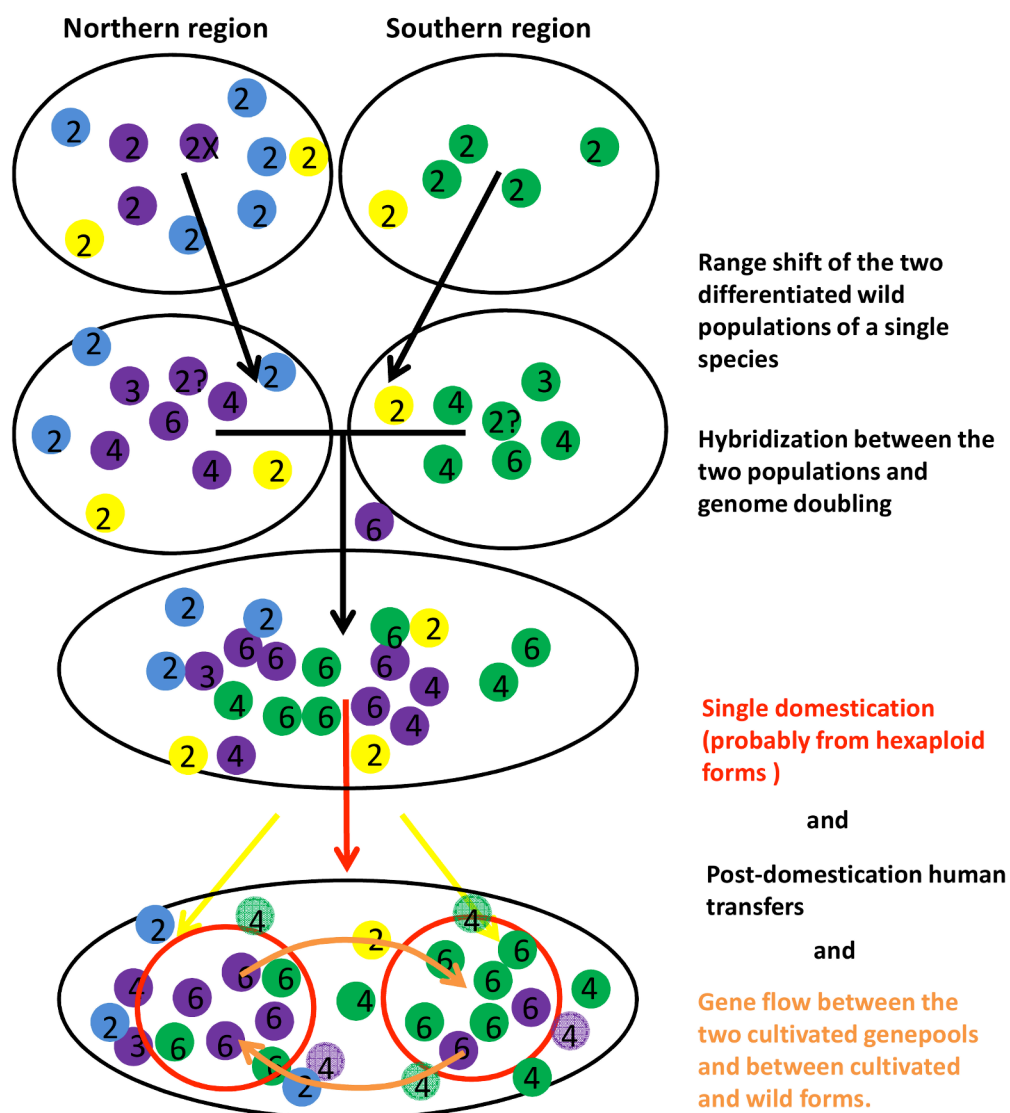
**Scenario 3:** Alternatively, this phylogeographical pattern might be related to:  
*Patterns of post-domestication human expansion*





**Scenario 4:** Cytological and marker-based studies suggest that the *I. batatas* hexaploid genome may be composed of two closely related genomes and a third one from a more distant relative [5,6]:

*Range shift of differentiated wild populations (climate-induced, for example) could have brought them into contact and allowed their hybridization. Intraspecific hybrids were then genetically stabilized by chromosome doubling. These range shifts may have concerned diploids, or individuals of intermediate ploidy levels (3X, or 4X and 6X). As hybridization often triggers genome structural and functional reorganization, which could explain the sudden appearance of new phenotypes, such hybridizations may have, in the case of sweet potato, favoured the appearance/stabilization of tuber-bearing mutants, the raw material for domestication. Patterns of post-domestication human expansion may finally have accentuated the clear-cut phylogeographical pattern found within cultivated *I. batatas*.*



The proposed scenarios are quite speculative at this stage. To provide more insights into where, how, and when polyploidization and sweet potato domestication occurred, marker-based studies should consider both homologous (similar gene copies within a progenitor diploid genome) and homeologous (similar gene copies brought together in a polyploid genome after the hybridization and genome doubling of differentiated diploid progenitors) copies of domestication genes, with a comprehensive sampling of both wild and cultivated *I. batatas*.

## References cited

1. Becerra Lopez-Lavalle LA, Orjeda G (2002) Occurrence and Cytological Mechanism of 2n Pollen Formation in a Tetraploid Accession of *Ipomoea batatas* (Sweet Potato). *J Hered*. 93: 185-192.
2. Freyre R, Iwanaga M, Orjeda G (1991) Use of *Ipomoea trifida* (HBK.) G. Don germ plasm for sweet potato improvement. 2. Fertility of synthetic hexaploids and triploids with 2n gametes of *I. trifida*, and their interspecific crossability with sweet potato. *Genome Res* 34: 209-214.
3. Orjeda G, Freyre R, Iwanaga M (1990) Production of 2n pollen in diploid *Ipomoea trifida*, a putative wild ancestor of sweet potato. *J Hered* 81: 462-467.
4. Parisod C, Holderegger R, Brochmann C (2010) Evolutionary consequences of autopolyploidy. *New Phytol* 186: 5-17.
5. Buteler MI, Jarret RL, LaBonte DR (1999) Sequence characterization of microsatellites in diploid and polyploid *Ipomoea*. *Theor Appl Genet* 99: 123-132.
6. Magoon ML, Krishnan R, Vijaya Bai K (1970) Cytological evidence on the origin of sweet potato. *Theor Appl Genet* 40: 360-366.

## *Chapitre 5*

*Article : Origine de la diversité de la patate douce  
(Ipomoea batatas (L.) Lam.) en Nouvelle  
Guinée, un second centre de diversité*

*On the origin of sweet potato (Ipomoea batatas  
(L.) Lam.) genetic diversity in New Guinea, a  
secondary centre of diversity*

**1. Title:** On the origin of sweet potato (*Ipomoea batatas* (L.) Lam.) genetic diversity in New Guinea, a secondary centre of diversity

**2. Authors' names:** C. Roullier<sup>1,2</sup>, R. Kambouo<sup>3</sup>, J. Paofa<sup>3</sup>, D. McKey<sup>2</sup> and V. Lebot<sup>1</sup>

**3. Full postal address:**

<sup>1</sup> CIRAD-BIOS, AGAP, Avenue Agropolis, 34398 Montpellier Cédex 5, France

<sup>2</sup> CEFE, UMR 5175, 1919 route de Mende, 34293 Montpellier Cédex 5, France ; and Institut Universitaire de France

<sup>3</sup> NARI, Southern regional centre, Central Province, Port Moresby, Papua New Guinea.

**4. Corresponding author:**

Roullier Caroline, CIRAD, Avenue Agropolis, 34398 Montpellier Cédex 5 France. CEFE, 1919 route de Mende, 34293 Montpellier Cédex 5, France; E-mail address: [caroline.roullier@cefe.cnrs.fr](mailto:caroline.roullier@cefe.cnrs.fr).

**5. Running title:** Sweet potato genetic diversity in New Guinea

**6. Article submitted to Heredity. Accepted with minor revisions.**

## ABSTRACT

New Guinea is considered the most important secondary centre of diversity for sweet potato (*Ipomoea batatas*). We analysed nuclear and chloroplast genetic diversity of 417 New Guinea sweet potato landraces, representing agro-morphological diversity collected throughout the island, and compared this diversity with that in tropical America. The molecular data reveal moderate diversity across all accessions analysed, lower than that found in tropical America. Nuclear data confirm previous results suggesting that New Guinea landraces are principally derived from the Northern neotropical genepool (Camote and Batata lines, from the Caribbean and Central America). However, chloroplast data suggest that South American clones (early Kumara line clones or, more probably, later re-introductions) were also introduced into New Guinea and then recombined with existing genotypes. The frequency distribution of pairwise distances between New Guinea landraces suggests that sexual reproduction, rather than somaclonal variation, has played a predominant role in the diversification of sweet potato. The frequent incorporation by farmers of plants issued from true seed, and the geographical and cultural barriers constraining crop diffusion in this topographically and linguistically heterogeneous island, have led to the accumulation of an impressive number of variants. As the diversification of sweet potato in New Guinea is primarily the result of farmers' management of the reproductive biology of their crop, we argue that on-farm conservation programs that implement distribution of core samples (clones representing the useful diversity of the species) and promote on-farm selection of locally adapted variants, may allow local communities to fashion relatively autonomous strategies for coping with ongoing global change.

**Keywords:** chloroplast and nuclear SSRs, secondary centre of diversity, germplasm conservation, sweet potato, Papua New Guinea

## INTRODUCTION

Crop genetic diversity is shaped by complex interactions between natural processes and those driven by humans. Diversity tends to be concentrated in particular parts of a crop's range (Pickersgill, 1998). Primary centres of diversity occur in areas where the plant was domesticated. Here, diversity is generated by long periods of interaction between humans and their plants, often with continued gene flow with the crop's wild relatives. Secondary centres of diversity occur in areas where the crop was introduced. In these areas, despite bottlenecks associated with introduction, the shorter period of cultivation and the usual absence of wild relatives, crops have produced huge diversity in a short time.

Our study focuses on the case of sweet potato (*Ipomoea batatas* (L.) Lam.) diversification in New Guinea. This large island is considered to be the most important secondary centre of genetic diversity for sweet potato, particularly the highlands region, where the total number of cultivars grown has been roughly estimated to be about 5000 (Yen, 1974; Bourke, 2009). Sweet potato is by far the most important crop in New Guinea. It strongly dominates agricultural production in the highland areas, and this was already the case when Europeans discovered the Central New Guinea highlands during the early 1900s. The dense human population encountered appeared to be heavily dependent on *I. batatas* for food and as fodder for pigs, and the crop played a central role in cultural rituals such as ceremonial exchange systems (Ballard, 2005). It has been argued that sweet potato triggered an "Ipomoean revolution" in this isolated region. Its adoption, largely replacing traditional crops such as taro (*Colocasia esculenta*), led to rapid demographic growth and transformed societies (Barrau, 1957; Yen, 1974; Ballard, 2005).

Sweet potato originated from tropical America. The nature of its introduction into New Guinea and the timing of its adoption—key points for understanding how its great diversity developed—are both unclear. According to archaeological, linguistic and historical data, sweet potato could have reached the Pacific by three different main paths (Barrau, 1957; Yen, 1974) corresponding to the introduction of clones from different geographical origins at different times. The *Kumara* line is based on the hypothesis of prehistoric introductions of clones by Polynesian sailors from coastal areas of South America to eastern Polynesia (Cook, Society and Marquesas Islands), around 1100 to 1200 AD, according to archaeological and linguistic evidence (Green, 2005). Sweet potato could have also reached Hawaii, New Zealand and Easter Island before early European contacts according to archaeological evidences (Ladefoged *et al.*, 2005). The *Batata* line assumes that later, around 1550 AD, Portuguese explorers could have transferred West Indian cultivars to Africa, India and their colonies in the Moluccas in eastern Indonesia. Finally, the *Camote* line supposes that Spanish "Acapulco-Manila" galleons could have spread sweet potato clones from Mexico to the Philippines very soon after Spanish conquest of these islands in 1522.

A recent molecular study (Roullier *et al.*, 2011) analysed 329 landraces collected from Mexico to Peru by nuclear and chloroplast microsatellite markers. Both kinds of markers supported the existence of two geographically restricted genepools, corresponding to accessions from the Peru-Ecuador region of South America (hereafter called the Southern genepool) and accessions from the Caribbean and Central America region (hereafter called the Northern genepool). According to the tripartite hypothesis of sweet potato introduction into Oceania<sup>2</sup>, the three lines introduced different genepools into different areas. Based on palynological and soil depositional evidence<sup>3</sup>, it has been suggested that sweet potato could have arrived in the New Guinea highlands in pre-Magellanic time (Golson, 1977; Gorecki, 1986; Haberle, 1998; Haberle and Atkin, 2005) via the *Kumara* line. Linguistic data have also been considered to support early introductions (Scaglione and Soto, 1994). In contrast, a recent review of the evidence suggests that sweet potato reached New Guinea only in the

---

<sup>1</sup> Oceania corresponds here to the entire insular region between Asia and the Americas, including Polynesia, Micronesia and Melanesia.

<sup>3</sup> In the absence of direct evidence for sweet potato in pollen records, palaeoecological studies have used sediment and pollen indicators of increased landscape clearance and degradation to infer its presence in New Guinea.

post-Magellanic period (around 1650), through local trade relations from the Moluccas to West New Guinea<sup>4</sup> (Ballard *et al.*, 2005). Based on surveys of oral history (Wiessner, 2005; Bourke and Harwood, 2009) and on archaeological data (Bayliss-Smith *et al.*, 2005), it has been proposed that sweet potato was introduced into the New Guinea highlands around 1700, where the plant was then relatively rapidly adopted and diffused. In the lowlands, the situation is different. It seems that although sweet potato may have been cultivated early in some parts of the northern coastal lowlands, introduction in most areas occurred after 1870 and its adoption as a major component of lowland agricultural systems has taken place since 1940, giving rise to a second “Ipomoean revolution” (Bourke, 2005). New Guinea thus appears to be a potential convergence point of these three lines, where both neotropical gene pools, Southern and Northern, could have been introduced at different times.

In the past decade, molecular markers have shown their potential to investigate origins of Pacific crops (Lebot, 1999; Kreike *et al.*, 2004; Clarke *et al.*, 2006; Hinkle, 2007; Gunn *et al.*, 2011; Perrier *et al.*, 2011), complementing archaeological and anthropological data. Molecular markers have already been used to analyse the genetic base of sweet potato from Oceania and New Guinea (Zhang *et al.*, 1998; Gichuki *et al.*, 2003; Zhang *et al.*, 2004) but studies thus far have used relatively small samples (fewer than 150 accessions). Their results show that varieties from Oceania are closely related to the Mexican accessions but only weakly related to those from Peru-Ecuador, suggesting that varieties from Oceania may be only of Mesoamerican origin.

Several sources of variation may contribute to the diversification of clonally propagated crops by farmers (McKey *et al.*, 2010). Somatic mutation is an important source of phenotypic variation in clonally propagated crops, especially in the Pacific (Lebot, 1992; Sardos *et al.*, 2008), and epigenetic inheritance may also contribute (McKey *et al.*, 2010). In some clonally propagated crops, sexual reproduction is certainly a key mechanism in the generation of diversity (McKey *et al.*, 2010). Variants selected by farmers are then maintained and multiplied by clonal propagation. Increasing evidence shows that many clonally propagated crops exhibit a mixed reproductive system in which evolutionary dynamics result from the interaction of clonal and sexual reproduction (Elias *et al.*, 2001; Caillon *et al.*, 2006; Scarcelli *et al.*, 2006; Sardos *et al.*, 2008; Delêtre, 2010). It has been reported that New Guinea farmers traditionally adopt sweet potato plants issued from true seed, resulting from the spontaneous germination of the numerous self-sown seeds the crop produces in New Guinean farming systems (Yen, 1974; Schneider, 1995). Incorporation of volunteer seedlings, selected by farmers across a wide array of ecological and cultural conditions throughout this large island, could be a major mechanism for creating diversity (Yen, 1974; Fajardo *et al.*, 2002).

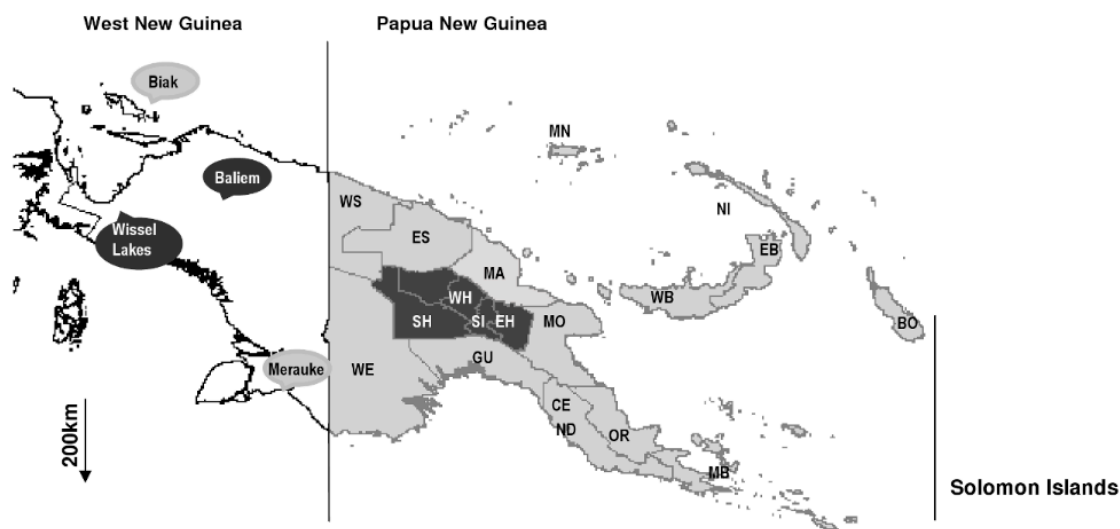
In Papua New Guinea, about 1000 accessions of sweet potato are maintained in *ex situ* collections at the Highlands Agricultural Experiment Station (HAES) at Aiyura in the Eastern Highlands Province and at the Lowlands Agricultural Experiment Station (LAES) at Keravat in East New Britain. Maintaining and evaluating such collections is laborious and expensive (Lebot *et al.*, 2005). A comprehensive study of New Guinea sweet potato genetic diversity is needed to eliminate duplicates and to identify the diverse sources of germplasm in the core collections (Brown, 1989). We analysed a subsample of accessions from these collections, representative of the agro-morphological diversity they contain. Using both chloroplast and nuclear microsatellites, we analysed the extent of diversity existing in subsets of both the HAES and LAES collections in order to characterize the distribution of allelic diversity throughout Papua New Guinea (samples available from West New Guinea were also included) and compare this diversity with that existing in tropical America. In particular, we analysed the origin of New Guinea sweet potato landraces in both highland and lowland agroecosystems. Finally, we attempted to explain the factors that contributed to the development of a secondary centre of genetic diversity in New Guinea.

---

3 Names for West New Guinea (the Indonesian part of New Guinea) are multiple and contested. Irian Jaya was generally used, but this name has recently been replaced by Papua. However, Papuans favouring independence refer instead to West Papua. We have chosen to call this region West New Guinea in this article.

**Figure 5-1: Geographic provenience of the 417 New Guinean sweet potato landraces from the NARI and NIAS collections, and definition of the regional and agro-ecosystem groups used in our study.**

The island of New Guinea is divided into West New Guinea and Papua New Guinea, the latter being divided into 20 provinces, each abbreviated by two letters on this map. Our subdivision for genetic analysis distinguishes sampled areas from the highlands in black (Wissel Lakes and Baliem localities for West New Guinea highlands and EH, SH, SI and WH provinces for Papua New Guinea highlands), to the lowlands in gray (Biak and Merauke localities for West New Guinea). Lowlands of Papua New Guinea were also subdivided into three regional groups: the south coasts of the mainland (WE, GU, CE and ND provinces), the north coast of the mainland (OR, MO, MA, ES and WS provinces) and the islands region (MB, BO, WB, EB, NI, and MN provinces).



Abbreviations used for Papua New Guinea Provinces: WE for Western, GU for Gulf, CE for Central, ND for National Central District, OR for Oro, MB for Milne Bay, MO for Morobe, MA for Madang, ES for East Sepik, WS for West Sepik, MN for Manus, NI for New Ireland, WB for West New Britain, EB for East New Britain, BO for the autonomous region of Bougainville, EH for Eastern Highland, WH for Western Highland, SH for Southern Highland, and SI for Simbu.



## MATERIALS AND METHODS

### *Plant materials and areas sampled*

A total of 369 sweet potato landraces was collected from the HAES and LAES collections of the National Agricultural Research Institute (NARI) of Papua New Guinea. Accessions were selected to represent agro-morphological variation across the range of sweet potato cultivation in highlands and lowlands of the country (Figure 5-1). We also analysed 48 accessions from the “Yen collection” maintained in The National Institute for Agrobiological Sciences (NIAS), in Tsukuba, Japan, including 30 accessions originating from Papua New Guinea, and 18 from West New Guinea (eight from lowland and 10 from highland regions). Healthy young leaves were collected from accessions maintained in greenhouses or fields and dried in an oven at 37 °C for two days, then conserved in silica gel. All DNA extractions (for both chloroplast and nuclear DNA) were conducted using the Qiagen 96 Plant kit for lyophilised tissues (Hilden, Germany), which proved suitable for tissues dried by our methods.

We recognized *a priori* two levels of hierarchical subdivisions for genetic analysis according to agro-geographical criteria (Figure 5-1). At the regional scale within Papua New Guinea, we distinguished the northern coast, the southern coast, the islands region and the highlands region, and within West New Guinea, the highlands and lowlands regions. Finally, we grouped these regions into New Guinea lowland and highland agroecosystem types (Figure 5-1), following a conventional subdivision in New Guinea agriculture (Bourke, 2009).

### *Chloroplast and nuclear genotyping and coding*

All individuals were genotyped with six chloroplast microsatellites (ccmp2, NTCP 26, NTCP 28, Ibcp 5, Ibcp 8 and Ibcp 10) and 11 nuclear microsatellites (J263, J315E, J522A, J116a, Ib297, J206A, J1809E, IbR16, IbC5, J544b, IbS11) described in a previous study (Roullier *et al.*, 2011). All loci were amplified independently using Multiplex PCR Taq (Qiagen) in a final volume of 10 µL, using 30ng of DNA per reaction. The following programme was conducted using a PTC-100 Thermocycler (MJ Research, Waltham, MA, USA): 15 min at 95°C, 35 cycles of 30 s at 94°C, 1 min 30 s at 57°C, 1 min at 72°C and finally 30 min at 72°C. Ninety-six-well plates were used for amplification and three wells per plate were devoted to negative (one well) and migration controls (two wells). The reproducibility of reactions was checked using replicated samples (25% of the total amount of reactions) on distinct plates; error rate was < 0.05 for the nuclear dataset and non-replicated samples were counted as missing data. For the chloroplast dataset, we only kept accessions with complete data. Amplification products were analysed with an ABI 3130 XL 16-capillary sequencer (ABI Prism, Applied Biosystems, Foster City, CA, USA) and allele scoring was eye-checked by two investigators using Genemapper (Applied Biosystems, Foster City, CA, USA).

### *Geographical patterns of nuclear and chloroplast genetic diversity in New Guinea landraces*

Spatial patterns of genetic diversity were inferred by computing different diversity estimators for both chloroplast and nuclear SSRs at different pre-defined geographical scales (regional and agroecosystem types [highlands or lowlands]). A rarefaction procedure was applied in some cases to account for unequal sampling among geographical areas (Petit *et al.*, 1998).

Nuclear SSR: As sweet potato varieties are clonally propagated, we estimated first the number of unique multilocus genotypes (MLGs) and then identified sets of clones or genets (pairwise genetic distances equal to 0) and clonal lineages (MLLs) by screening each MLG pair presenting extremely low genetic distances (here individuals differing in fewer than six variables). These pairs accounted for a small peak (at very low values) in the frequency distribution of distances, making this distribution bimodal rather than unimodal (Arnaud-Haond *et al.*, 2007). Two indices of clonal diversity, *Gc* and *Gl* (number of MLGs or of MLLs, respectively, divided by the total number of accessions), were calculated. The R software was used to obtain the frequency distribution of pairwise distances (Manhattan distance) (R Development Core Team, 2009) within New Guinea lowland and highland groups separately. We then computed the mean number of alleles (*Na*), the allelic richness (*Ar*, the rarefied mean number of alleles per locus), and the rarity index, *R*. *R* is a rarefied index that

**Table 5-1: Nuclear and chloroplast diversity within regional and agroecosystem-type groups in New Guinea.**

<b>Regions</b>	<i>N</i>	<i>H</i>	<i>H1_2</i>	<i>H6_9</i>	<i>H14</i>	<i>H13</i>	<i>H12</i>	<i>H5</i>	<i>DH</i>	<i>NA</i>	<i>Ar</i>	<i>R</i>
<i>Northern coast of Papua New Guinea</i>	40	4	21	3	13		1		0.552 (0.057)	7.273	5.933 (0.660)	0.230 (0.050)
<i>Islands of Papua New Guinea</i>	118	4	73	14	28	3			0.542 (0.041)	8.273	5.942 (0.780)	0.260 (0.053)
<i>Southern coast of Papua New Guinea</i>	12	3	8	1	2				0.473 (0.162)	5.910	5.655 (0.326)	0.179 (0.012)
<i>West New Guinea lowlands</i>	8	2	2		6				0.476	5.022	5.022	0.235
<b>Total lowlands</b>	179	5	104	18	49	3	1		0.510 (0.086)	8.546	8.500 (0.158)	0.255 (0.004)
<i>Highlands of Papua New Guinea</i>	229	6	109	22	83	9	3	1	0.687 (0.009)	8.728	5.701 (0.790)	0.215 (0.053)
<i>West New Guinea highlands</i>	10	2	6		2				0.500 (0.128)	4.900	4.845 (0.160)	0.260 (0.019)
<b>Total highlands</b>	238	6	115	22	85	9	3	1	0.595 (0.068)	8.73	8.349 (0.448)	0.211 (0.006)

All diversity indices were computed following a rarefaction procedure (by re-sampling eight and 150 individuals (1000 times each) for comparison of regional groups and agroecosystem groups, respectively). *N* is the total number of individuals sampled, *H* the number of distinct haplotypes, *DH* the rarefied haplotype diversity, *NA* the mean number of alleles, *Ar* the allelic richness and *R* the rarity index. Standard deviations are indicated between brackets.

reflects the proportion of rare alleles in a given dataset, being inversely proportional to the allele frequency in the global dataset.  $Ar$  and  $R$  were averaged from 1000 resamplings of eight individuals for the regional scale and 150 for the agroecosystem-type scale. Number of individuals for re-sampling was set by the smallest number of individuals found in one of the six regions and in one of the two agroecosystem types, respectively. Computations were made using custom R scripts.

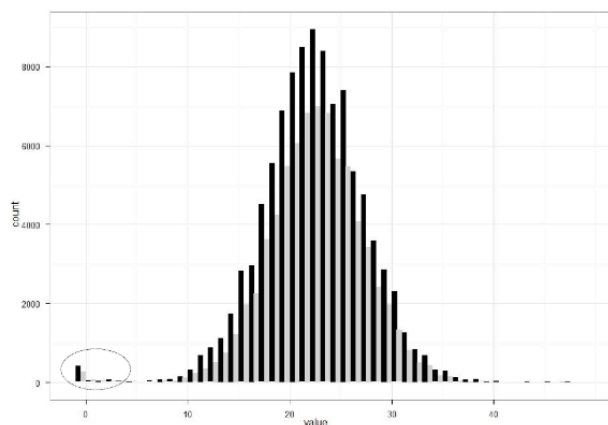
The partitioning of genetic variation among agroecosystem types and regional groups was quantified with an AMOVA using Arlequin software 3.1 (Excoffier *et al.*, 2005). We also computed a Principal Coordinate Analysis (PCoA) based on a “Lynch” distance matrix between individuals (“polysat” R CRAN package; Clark and Jasieniuk, 2011). We then performed a one-way ANOVA between PCoA first axis coordinates and agroecosystem-type groups, followed by a Tukey test to assess the significance of the differentiation following this axis (the most discriminating for the dataset) between the New Guinea highlands and lowlands groups, using R software (aov and TukeyHSD functions, R Development Core Team, 2009).

**Chloroplast SSRs:** We identified distinct haplotypes and plotted their geographic distributions. We computed  $H_r$ , the rarefied haplotypic diversity, for regional groups using the program CONTRIB (Petit *et al.*, 1998) (results averaged from re-samplings of eight individuals).

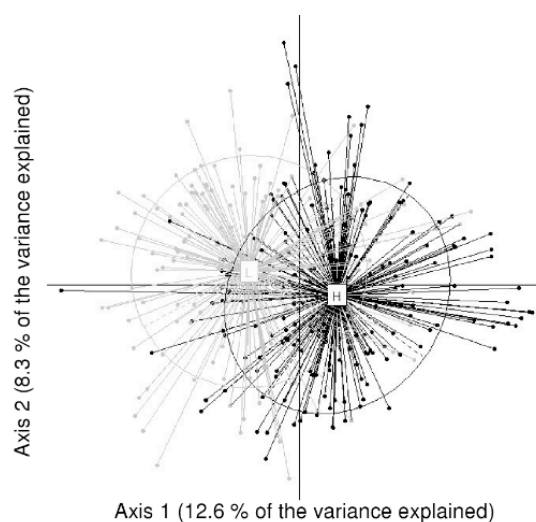
**Figure 5-2: Patterns of genetic differentiation of sweet potato landraces in New Guinea.**

a) Frequency distribution of genetic dissimilarity based on the calculation of all pairwise Manhattan distances between genotypes within landraces of New Guinea lowlands in gray, and New Guinea highlands in black. The first peak of the bimodal curve represents clones (pairwise dissimilarities equal to 0), possible variants by somatic mutations and/or scoring errors. b) Principal Coordinate Analysis based on the Lynch distance between samples. Highland and lowland genotypes appear in black and gray, respectively.

a)



b)



### ***Assessing the genetic relationships between New Guinean and tropical American genepools***

For purposes of comparison, we added to the New Guinea dataset nuclear and chloroplast datasets already published in a previous study (Roullier *et al.*, 2011). The tropical America dataset comprises 130 individuals (76 accessions from the Northern genepool and 54 from the Southern genepool) for nuclear data, and 329 individuals (192 from the Northern genepool and 137 from the Southern genepool) for chloroplast data.

**Comparison of chloroplast data:** We first compared haplotypes identified in New Guinea and tropical America. Our previous study showed that tropical American sweet potato landraces are partitioned in two geographically restricted cp lineages, cp 1 and 2, corresponding to most of the accessions from Southern and Northern genepools, respectively. We estimated the proportions of haplotypes from each lineage, in each province of Papua New Guinea and in collecting sites in West New Guinea.

**Comparison of nuclear data:** We aimed to identify the origin of New Guinean landraces and the relative contribution of Northern and Southern genepools from the neotropics to New Guinea genetic diversity. To do this, we attempted to “assign” New Guinean individuals to their potential sources, the Northern and Southern genepools, using two complementary kinds of assignment methods. First, we used a “general method”, the Discriminant Analysis of Principal Components (DAPC), a multivariate analysis recently developed and implemented in the adegenet R packages (Jombart, 2008) (Supplementary material 1). DAPC provides an efficient description of genetic clusters using a few synthetic variables (called the discriminant functions). This method seeks linear combinations of the original variables (alleles) which show differences between groups as best as possible while minimizing variation within clusters. Based on the retained discriminant functions, the analysis derives probabilities for each individual of membership in each of the different groups. This coefficient can be interpreted as “genetic proximity” of individuals to the different clusters and might provide an “assignment measure” of individuals to predefined groups.

In our case, we first constructed the linear model and obtained synthetic variables on the tropical America dataset. Then, we used New Guinean landraces as supplementary individuals (that were not used in constructing the model). We did this by transforming the New Guinea dataset using the “centring” and “scaling” of the Tropical America dataset, and then using the same discriminant coefficients as for the individuals that contributed to model construction to predict the position of the new individuals on the discriminant functions. The analysis provides probabilities for each New Guinea accession that it originated from each of a number of predefined neotropical groups. DAPC in itself requires construction of prior groups. Then, we first ran a sequential K-means clustering algorithm for  $K = 2$  to  $K = 10$ , after transforming the data using a PCA (notably to reduce the number of variables and thereby speed the clustering algorithm). This allowed identifying an optimal number of genetic clusters to describe the data, by comparing the different clustering solutions using the Bayesian Information Criterion (BIC). Following this analysis, three genetic clusters were considered optimal to summarize the data. However, we consider grouping obtained for  $K = 2$ , an accurate and simpler summary of the tropical American genepools (see Supplementary material 1 for details of the DAPC analysis). We then performed DAPC for  $K = 2$ , retaining five PCA components (29.3% of the total variance) for prior data transformation, corresponding to the “optimal” value following the a-score optimization procedure proposed in adegenet (Jombart *et al.*, 2010). A landrace is considered to be well assigned to its source group if the associated membership probability is greater than 0.8.

We also compared non model-based DAPC assignment with that obtained by a Bayesian model-based method implemented in the software Structure 2.3.3 (Falush *et al.*, 2007). By pre-specifying source genepools in tropical America (as defined by the K-means clustering algorithm for  $K = 2$ ), the algorithm implemented in Structure estimates ancestry for additional individuals (from New Guinea), updating allele frequencies using only those from tropical America. We ran the admixture model with  $K = 2$ , correlated allele frequencies, 50 000 burn-in iterations and 150 000 Markov chain-Monte Carlo steps and data coding for handling genotype ambiguity for co-dominant markers in polyploids (Falush *et al.*, 2007).

Finally, we performed a Principal Coordinate Analysis (PCoA) computed on a “Lynch” distance matrix between tropical American and New Guinean accessions to describe genetic variation between New Guinean landraces and tropical American genepools and to quantify the partitioning of

**Table 5-2: Analysis of molecular variance (AMOVA) of nuclear SSRs within New Guinea (among agroecosystem groups, among regional groups and within regional groups) and among tropical America (Southern and Northern genepools) and New Guinea.**

Source of variation	Sum of squares	Variance components	Variation percentage	Fixation indices
<b>Within New Guinea</b>				
<i>Among agroecosystem groups</i>	97.600	0.119	1	FCT = 0.010
<i>Among regional groups</i>	90.259	0.424	3.590	FSC = 0.036
<i>Within regional groups</i>	4640.420	11.290	95.410	FST = 0.046
<b>Among tropical America and New Guinea</b>				
<i>Among groups "New Guinea-neotropical Northern"</i>	176.139	1.27855	9.810	FST =0.098
<i>Within groups</i>	5773.581	11.759	90.190	
<i>Among groups "New Guinea-neotropical Southern"</i>	236.245	2.350	16.930	FST= 0.169
<i>Within groups</i>	5407.241	11.52930	83.07	

genetic variation among tropical American (total, neotropical Southern and Northern genepools) and New Guinean accessions with an AMOVA using Arlequin software. We also measured differentiation between groups, calculating intra- and inter-group mean pairwise distances (Lynch distance, R package polysat) between genotypes and a measure of  $F_{ST}$  based on the estimation of allelic frequencies (“simpleFreq” and “calcFst” functions of the R package polysat).

To compare allelic diversity between the neotropics and New Guinea, we computed  $A_r$ ,  $R$ ,  $H_e$  (expected heterozygosity, derived from allelic frequencies estimated using the “simplefreq” function of the R package polysat),  $H_o$  (observed heterozygosity),  $R_a$  (the allelic size range),  $G_c$  and  $G_l$  (indices of clonal diversity) and  $H$  (total number of haplotypes) using custom R scripts. All indices were estimated following a rarefaction procedure (except for  $NA$ ,  $G_c$ ,  $G_l$  and  $H$ ), where results were averaged from the re-sampling of 50 individuals for the comparison of Northern and Southern neotropical genepools and New Guinea highlands and lowlands groups, and 130 for the global comparison of New Guinean and tropical American accessions. In order to explain observed neutral diversity calculated in terms of allelic richness, we also established a curve characterising the increase in the total number of captured alleles for a given sample size (based on a 1000-re-samplings procedure) for each group. This curve allowed us to estimate the number of clones that have been introduced from tropical America to New Guinea.

**Table 5-3: Comparison of nuclear and chloroplast diversity between tropical America and New Guinea.**

Landraces groups	Regions	<i>Nnuc</i>	<i>NA</i>	<i>Ar</i>	<i>He</i>	<i>Ho</i>	<i>ra</i>	<i>R</i>	<i>K1</i>	Private alleles (>0.05)	<i>Gc</i>	<i>Gl</i>	<i>Ncp</i>	<i>cpDNA 1</i>	<i>H</i>
Tropical America	<b>Total</b>	130	10.91	10.9	0.769	0.367	25.455	0.354	-	24 (7)	0.94	0.83	329	-	21
	<i>Northern</i>	76	9.727	9.35	0.768	0.385	24.57	0.371	0.276	9 (3)	0.99	0.8	192	14.7 (37.9)	13
	<i>Southern</i>	54	8.455	8.372	0.733	0.343	24.247	0.332	0.724	8 (2)	1	0.87	137	73.9 (74.4)	15
New Guinea	<b>Total</b>	417	9.545	8.3	0.747	0.387	23.074	0.169	0.716	16 (3)	0.9	0.79	407	36.27	6
	<i>Highlands</i>	239	8.818	7.381	0.734	0.381	22.153	0.158	0.264	6 (0)	0.9	0.727	236	41.94	6
	<i>Lowlands</i>	178	8.636	7.63	0.758	0.396	23.863	0.186	0.310	4 (0)	0.95	0.872	171	28.48	6

*Nnuc* is the total number of accessions for the nuclear dataset, *NA* the mean number of alleles. Five indices (averaged over all loci) were computed following a rarefaction procedure based on 1000 re-samplings of 130 accessions for New Guinea and tropical America total datasets and 50 accessions each for neotropical Southern, neotropical Northern, New Guinea highland and New Guinea lowland datasets. *Ar* is allelic richness, *He* the expected heterozygosity, *Ho* the observed heterozygosity, *ra* the allelic size range, *R* the rarity index and *K1* the mean *K1* ancestry value given by the Bayesian clustering analysis. *Gc* and *Gl* are indices of clonality (giving the ratio of the number of unique multilocus genotypes and multilocus clonal lineages, respectively, identified from nuclear data to the number of samples). *Ncp* is the total number of accessions of the chloroplast dataset, *cpDNA 1* the frequency of *cpDNA* lineage 1 haplotypes and between brackets the same index considering only the nuclear sub-sampling, and *H*, the number of distinct haplotype.



## RESULTS

### *Geographical patterns of genetic diversity in New Guinea*

**Chloroplast markers:** We identified a total of six different haplotypes in our sample of New Guinea landraces (H1\_2, H5, H6\_9, H12, H13, H14). Haplotype richness varied from 0.473 ( $\pm 0.162$ ) in the southern coastal region to 0.687 ( $\pm 0.009$ ) in the highlands of Papua New Guinea (Figure 5-1). H5 was found only once, in the highlands of Papua New Guinea. H12 was found in both New Guinea highland and lowland sites but was very rare (freq  $< 0.01$ ). Another uncommon haplotype, H6\_9 (freq = 0.091), was also found in both agroecosystem groups, while haplotype H13 (freq = 0.027) was present mostly in the New Guinea highlands. H14 and H1\_2 were clearly dominant (freq = 0.33 and 0.5, respectively) and were distributed over all areas sampled (Table 5-1).

**Nuclear markers:** We found a total of 378 multilocus genotypes and 329 putative clonal lineages (Manhattan distance between genotypes within a lineage  $< 6$ ) ( $Gc = 0.9$  and  $Gl = 0.79$ ). Duplicates were sometimes locally restricted but were also shared between distant regions (shared notably between New Guinea highland and lowland regions). Frequency distribution of pairwise Manhattan distances for New Guinea highland and lowland genotypes showed a bimodal curve, attesting to the existence, in both groups, of clones (identical MLGs) and of apparent clonal lineages (or sampling errors) differing by only few somatic mutations. Both occurred at low frequency ( $Gc = 0.9$  and  $0.94$  and  $Gl = 0.72$  and  $0.87$  in New Guinea highlands and lowlands, respectively) (Figure 5-2a).

The mean number of alleles per locus ranged from 4.9 in the highlands of West New Guinea to 8.721 in the highlands of Papua New Guinea, with a mean value for all six regions of 6.68 (Table 5-1). Values of allelic richness, an estimator that allows comparison of samples of unequal size, ranged from 4.845 ( $\pm 0.160$ ) in the highlands of West New Guinea to 5.942 ( $\pm 0.78$ ) in the islands region of Papua New Guinea. Similar values for allelic richness were found in New Guinea highland and lowland landraces ( $8.349 \pm 0.448$  and  $8.5 \pm 0.158$ ), respectively). Values for the rarity index ranged from 0.179 ( $\pm 0.012$ ) in the southern coastal region of Papua New Guinea to 0.26 for the islands region of Papua New Guinea ( $\pm 0.053$ ) and highlands of West New Guinea ( $\pm 0.019$ ). Values of the rarity index in New Guinea highland and lowland agroecosystems were 0.211 ( $\pm 0.006$ ) and 0.255 ( $\pm 0.004$ ), respectively.

Results of the AMOVA underlined the weak genetic differentiation between New Guinea highland and lowland agroecosystems and among regions (Table 5-2). Most of the variation occurred within regional groups (95.5 %), with little variation between highland and lowland systems and among regions (1 and 3.5 %, respectively). Results of the PCoA confirmed the weak differentiation among regions (Supplementary material 2) and between ecosystem types within New Guinea sweet potato landraces. However, the first axis, which accounted for 12.6 % of the variance, still showed a slight but significant ( $p < 0.01$ ) differentiation between New Guinea lowland and highland accessions (Figure 5-2b).

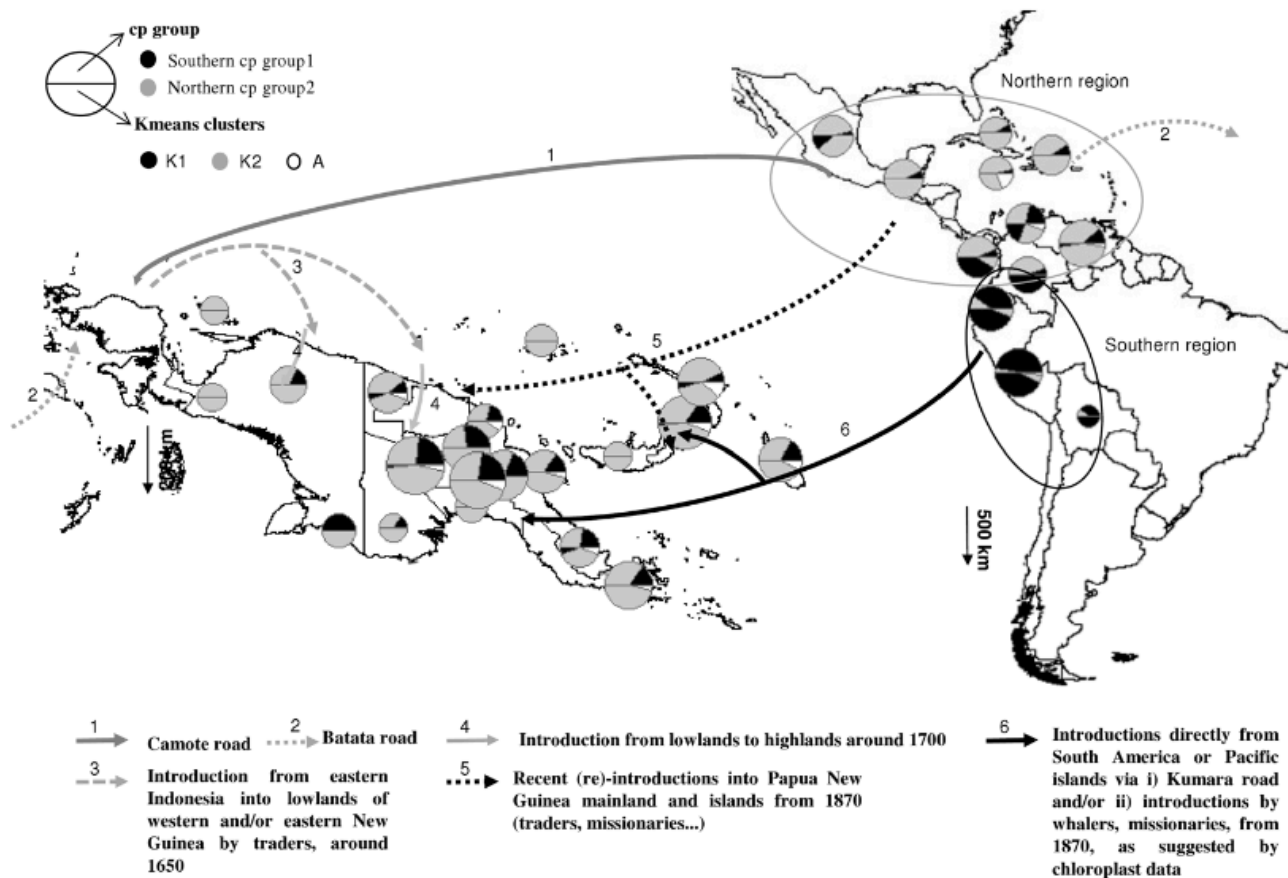
### *Comparing genetic diversity between tropical American and New Guinean landraces*

**Chloroplast markers:** Chloroplast diversity found in New Guinea was a subset of that present in tropical America (Table 5-3). Out of 21 haplotypes found in tropical America, six were found in New Guinea (H1\_2, H5, H6\_9, H12, H13, H14). New Guinea haplotypes included representatives of both cp groups 1 and 2, defined in a previous study (Roullier *et al.*, 2011): H14 and H13 were the most frequent haplotypes of cp group1. Widespread throughout the Southern region in the neotropics, they are also present in the Northern region, although H13 is rare there. H6\_9 and H1\_2 are the most frequent haplotypes of cp group 2, found throughout the Northern region, but are also present in the neotropical Southern genepool. H5 and H12 are very rare cp group1 haplotypes, restricted to Peru and Bolivia, respectively (Supplementary material 3).

Both cp groups 1 and 2 were represented throughout New Guinea, but cp group 2 dominated (63% on average), except for the highlands of Papua New Guinea, where the proportion of cp group 1 haplotypes was 42% (Table 5-1 and Figure 5-3). Minor cp group 1 haplotypes (H5, H12 and H13) were found more frequently in New Guinea highland than in New Guinea lowland sites.

**Figure 5-3: Geographical origin of New Guinea and Tropical America sweet potato landraces (as distinguished by chloroplast data, and nuclear data with DAPC analysis at  $K = 2$ ) and possible paths of sweet potato introduction into New Guinea.**

Size of circles is proportional to the total number of individuals by country, province and sampled area, for tropical America, Papua New Guinea and West New Guinea respectively. In the pie charts, the top half of the circle represents the proportions of individuals belonging to each chloroplast group (cp group 1 and cp group 2) while the bottom half represent the proportions belonging to nuclear clusters K1 and K2. A means “Admixed” for individuals with a membership probability (DAPC) or ancestry value (Bayesian clustering) inferior to 0.8 for both clusters. Arrows indicate proposed routes and dates of introduction by humans.



Nuclear markers: For almost all of the estimators (*NA*, *Ar*, *Ra*, *R*, and *He*), nuclear genetic diversity tended to be greater in tropical America than in New Guinea (Table 5-3), particularly comparing that of the Northern genepool with that of New Guinea group. Only *Ho* was slightly higher in New Guinea (0.387), close to the value for this index (0.385) for the Northern group. Allelic richness of the Southern genepool (8.372) was equivalent to that found in New Guinea (8.3) but the rarity index was higher in the Southern genepool ( $R = 0.32$ ) than in New Guinea ( $R = 0.169$ ). The tropical America group had 24 private alleles and the New Guinea group 16 (all the latter were rare, with frequencies  $< 0.1$ ). Neotropical Northern and Southern genepools, and New Guinea lowland and highland groups, had 9, 8, 6 and 4 private alleles, respectively. Clonality indices *Gc* and *Gl* were respectively 0.90 and 0.79 for New Guinea accessions and 0.94 and 0.83 for those from tropical America.

We used grouping obtained with K-means clustering for  $K = 2$  to perform the DAPC and the Bayesian assignment analysis (supplementary material 1): the cluster *K1* mainly contained samples from the Southern genepool (79%), and cluster *K2* mostly those from the Northern genepool (91%) (Figure 5-4a). DAPC results showed that most of the New Guinea accessions (370 accessions), from both lowland and highland agroecosystem types, were clearly assigned (with a membership probability  $> 0.8$ ) to the cluster *K2* (mean *K2* membership probability of  $0.92 \pm 0.20$ ) (Figure 5-4a and 5-4b). Only few accessions (10) were found to be assigned to the cluster *K1* (mean *K1* membership probability of 0.08). Some others were not well assigned to a given cluster (membership probability  $< 0.8$ ) and are considered to be “admixed”. These included 37 accessions of uncertain assignment to genetic clusters *K1* or *K2*. Bayesian “assignment” provided congruent results (Figure 5-4a). New Guinea accessions exhibited a mean *K2* ancestry value of  $0.716 (\pm 0.116)$ , a value quite similar to that found in the Northern genepool ( $0.724 \pm 0.145$ ), while the Southern genepool exhibited a higher mean *K1* ancestry value ( $0.710 \pm 0.185$ ) (Table 5-3).

The PCoA analysis differentiated neotropical Northern and Southern neotropical genepools and the New Guinea samples, principally on the first axis (16.3% of variance explained) (Figure 5-4b). The Southern genepool was the most divergent from the New Guinea genepool, as also attested by highest values of both inter-group mean pairwise distance and *Fst* (Table 5-4), confirming DAPC results. Results of AMOVA also supported this tendency: 9.81 % of the total variance was between Northern and New Guinea genepools and 16.93% was between Southern and New Guinea genepools (Table 5-2). Moreover, PCoA showed that lowland accessions appear less divergent from the Northern genepool than highland accessions (Figure 5-4b).

**Table 5-4: Distance and differentiation between accessions from New Guinea (highland and lowland groups) and the neotropics (Northern and Southern genepools), measured by intra- and inter-group pairwise distances (Lynch distances) above the diagonal and Fst values below the diagonal.**

		New Guinea			neotropical	
		<b>Total</b>	<i>lowlands</i>	<i>highlands</i>	<i>Northern</i>	<i>Southern</i>
<b>New Guinea</b>	<b>Total</b>	0.341	-	-	0.403	0.434
	<i>lowlands</i>	-	0.343	0.349	0.405	0.433
	<i>highlands</i>	-	0.004	0.339	0.402	0.433
<b>neotropical</b>	<i>Northern</i>	0.006	0.008	0.010	0.386	0.433
	<i>Southern</i>	0.012	0.020	0.020	0.020	0.356

## DISCUSSION

### *History of sweet potato introduction into New Guinea: the relative contributions of Kumara, Batata and Camote lines*

The timing and nature of sweet potato introduction into New Guinea are still controversial. New Guinea forms the point at which the three lines, *Kumara*, *Batata* and *Camote*, could have converged and their relative contributions to New Guinea sweet potato diversity over time have been widely debated among archaeologists and anthropologists of Oceania. Some consider that the high degree of dependence between sweet potato and highlands societies is best explained as the result of a long period of evolution (implying ancient introduction via the *Kumara* line). Opponents argue that widespread adoption and diversification could have taken place within 300 to 400 years, a hypothesis compatible with 16<sup>th</sup>-Century transfers to New Guinea via the *Batata* and *Camote* lines. According to the proponents of this hypothesis, the intrinsic properties of sweet potato (its broad tolerance of environmental conditions and high productivity, its relative ease of cultivation), combined with innovations developed by highland societies in a constraining and heterogeneous environment, are sufficient to explain the intensity and relative rapidity of the adoption of sweet potato and its diversification in New Guinea (Ballard, 2005).

All results on nuclear markers show that New Guinean landraces are clearly divergent from the genotypes of the Southern genepool, while they are only weakly differentiated from those of the Northern genepool. These results confirm previous results (Gichuki *et al.*, 2003; Zhang *et al.*, 2004) suggesting a main Central and Caribbean American origin for New Guinean landraces via the *Camote* and *Batata* lines (Figure 5-3).

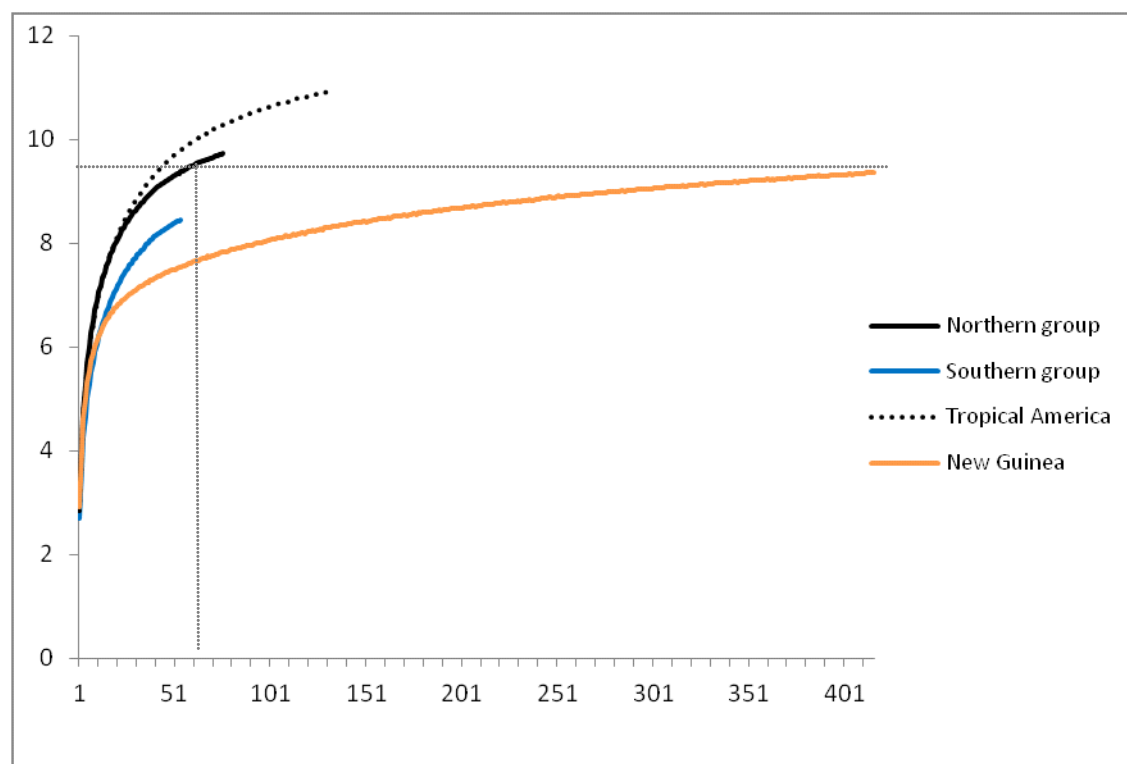
However, both cp groups (1 and 2) typical of the Southern and Northern genepools, respectively, were identified in New Guinea landraces. Two hypotheses could explain the presence of haplotypes typical of the Southern genepool in New Guinea (Figure 5-3). First, they may already have been present in clones introduced from the Northern genepool. Recombination between migrant clones of Northern and Southern genepools has occurred since ancient times in tropical America (Roullier *et al.*, 2011), blurring the phylogeographic pattern in chloroplast DNA diversity and complicating assignation of clones introduced into New Guinea to areas of origin in the Neotropics. In fact, the most common haplotypes of cp group 1 are also found, albeit at lower frequency (14.7%), in the Northern genepool. Second, clones of Southern origin may have been introduced into New Guinea, either directly from South America or from other regions in Oceania into which they had been introduced previously. Interestingly, three rare haplotypes of cp group 1, mostly restricted to the Southern genepool, were found in New Guinea, which could be seen as an indication of an early introduction through the *Kumara* Line.

Based on paleoecological data, some archaeologists have associated the intensification of land use, assessed from *Casuarina* pollen records at about 1200 BP, to an early introduction of sweet potato from Polynesia (Golson, 1977; Gorecki, 1986; Haberle, 1998). Following the same reasoning, linguists have argued that the presence of terms related to the Polynesian word “*kumara*” could support a hypothesis of early introduction (Scaglione and Soto, 1994). While our chloroplast data are compatible with the existence of such introductions, our nuclear data indicate that they cannot account for the main genetic background of New Guinean sweet potato diversity, either in the New Guinea highlands or lowlands. Such introductions may rather have been occasional, resulting more probably from post-Magellanic introductions of clones by whalers, missionaries and Polynesians and through 20<sup>th</sup>-Century movements across the Pacific (Bourke, 2005) (Figure 5-3). Alternatively, primary *Kumara* line introductions could have been widely reshuffled by later European re-introductions from the Northern genepool.

Our nuclear data also suggest a slight differentiation between landraces from New Guinea highlands and lowlands. Moreover, the highlands group is more differentiated from the ancestral neotropical Northern genepool than is the lowland group. Two complementary scenarios could explain this pattern. The observed differentiation could be largely drift-driven, resulting from a sequence in which sweet potato first colonized the island region, then mainland coastal areas, and finally the highlands. The slight decrease of genetic diversity from lowland to highland areas might be an

**Figure 5-5: Allelic richness that can be captured for a given number of samples.**

The minimum number (63) of clones introduced from the neotropical Northern genepool that is sufficient to explain allelic richness encountered in New Guinean sweet potato landraces is deduced from this curve.



indication of this human-mediated dispersal. Based on surveys of oral history, it has been proposed that sweet potato was introduced around 1700 into the New Guinea highlands, where the plant was rapidly adopted and diffused (Bourke, 2009). It is likely that sweet potato was available on the coasts of New Guinea as early as 1521 AD and certainly at least by 1633 AD. Therefore, sweet potato probably reached the highlands from the Sepik area along one of the corridors that link coastal regions to the Border Mountains. Furthermore, it seems that even though sweet potato was cultivated early in some lowland parts of New Guinea (north coastal areas of the mainland), its adoption in lowland regions dates mostly to a (re-)introduction after 1870 by Europeans, Polynesians and other outsiders. Widespread adoption of sweet potato into lowland agricultural systems has occurred since 1940 (Bourke, 2005). Thus, the weak differentiation between the islands region and the neotropical Northern genepool could be associated with repeated recent introductions, and thus a shorter time during which drift could have occurred (Figure 5-3).

### ***Processes of diversification: on the origin of a secondary centre of diversity***

Several examples of secondary diversification have been well documented: *Citrus* spp. in Europe (Ollitrault and Luro, 2001); watermelon, *Citrullus lanatus* [Thunb.] Mats. & Nakai in Brazil (Romão, 2000); bean, *Phaseolus vulgaris* L., in Africa (Pickersgill, 1998; Asfaw et al., 2009); barley, *Hordeum vulgare* L., in Africa (Pickersgill, 1998); and cassava, *Manihot esculenta* Crantz subsp. *esculenta*, in Africa (Pickersgill, 1998; Delêtre, 2010) and Vanuatu (Sardos et al., 2008). Causes of secondary diversification have been discussed (Pickersgill, 1998; Romão, 2002; Sardos et al., 2008; Delêtre, 2010). Multiple introductions of a crop can mitigate any bottleneck effects, especially if introductions have led to the mixing of different genepools from the original range. Subsequent recombination between genepools, associated with new selective pressures, both natural and cultural, could lead to the rapid selection of a large number of new variants.

Chloroplast markers indicate a reduction of diversity in New Guinea compared to tropical America (from 21 haplotypes to 6). The most frequent haplotypes from tropical America were found (H14, H1\_2, H6\_9) in New Guinea. Nuclear data also revealed that New Guinea accessions exhibited lower diversity than that observed for the original Northern genepool, and a level of diversity similar to that of the Southern genepool, except for rare alleles. These results highlight the introduction bottleneck. However, as the reduction of genetic diversity is relatively moderate, it suggests a quite limited bottleneck. Multiple independent introductions can reduce bottleneck effects by increasing the genetic diversity introduced into the new area. Such scenarios have been invoked to explain the impressive diversity of beans in East Africa (Pickersgill, 1998; Asfaw et al., 2009) and of watermelons in Brazil (Romão, 2000). Beans were domesticated independently in Mesoamerica and in the Andean region, giving rise to two differentiated genepools. Both were introduced into Africa, leading to a concentration of diversity at the scale of individual fields (Asfaw et al., 2009). Bean evolutionary history shows parallels with that of sweet potato. But as bean genotypes from both American genepools rarely recombined (Asfaw et al., 2009; Burle et al., 2010), it still remains easy to identify original sources in introduced areas.

This identification is less evident in our case, because we are dealing with a situation of initial admixture in tropical America before introduction into New Guinea, as detected by both nuclear and chloroplast data (Figure 5-3, Supplementary material 1; Roullier et al., 2011). We propose therefore that independent introductions of “pure” individuals from the Northern neotropical genepool (accessions with cp group 2), introgressed individuals (Northern genepool accessions with cp group 1), and probably a few individuals directly from the Southern neotropical genepool (Southern genepool accessions with cp group 1) could have led to accumulation of a substantial subset of the genetic diversity present in tropical America (Figure 5-4). We estimated that the introduction of 63 clones from the Northern genepool (or of 45 clones if these came from throughout tropical America) would have been sufficient to explain the current New Guinea diversity in terms of allelic richness (Figure 5-6).

Our data show, moreover, that few clonemates were identified among landraces through the island of New Guinea. The 417 landraces considered included 329 genotypes that were most likely derived from distinct sexual reproduction events (208 over 230 highland landraces, and 170 over 179 lowland landraces). The greatest proportion of the diversity among cultivars in New Guinea thus

appears to have resulted from intensive recombination among introductions, both in highland and in lowland conditions. This situation of secondary admixture may explain the difference observed between chloroplast and nuclear genetic signals: Introductions from the Southern genepool recombined with local material, most of which was of Central American and Caribbean origin (i.e., Northern genepool), blurring their original nuclear signal.

Sweet potato is vegetatively propagated by stem cuttings. Various studies have reported the widespread occurrence and use by farmers of plants issued from true seed in different New Guinea highland farming systems (Bulmer, 1965, Schneider, 1995; Yen, 1974). Cultivation practices favour the appearance and subsequent adoption of such variants, similarly to what has been described in detail and reviewed for cassava in the Neotropics (McKey *et al.*, 2010) and for taro in Oceania (Caillon *et al.*, 2006, Sardos *et al.*, 2011). In the highlands, sweet potato plants are harvested over a long period of time and then the parcel is left as fallow, giving the remaining plants ample time to reach flowering and fruiting stages. Moreover, flowering of sweet potato is recognized to be particularly abundant under highland conditions (P. Van Wijmeersch, pers. comm., 2011). Furthermore, mixtures of different varieties (landraces) are a ubiquitous characteristic of highland gardens. The number of varieties grown by different communities varies from 6 to 71, with a mean of 33 varieties per garden (Bourke, 2009), favouring even more genetic mixing. As sweet potato is an obligatorily outcrossing plant, with a sporophytic incompatibility system, seeds are produced only if different incompatibility groups are planted in the same garden (or in gardens close enough to permit pollen flow). Seventeen auto-incompatibility groups have been described in tropical America (Nakanishi and Kobayashi, 1979). Multiple introductions may have facilitated the arrival of several incompatibility groups, leading to high probability of effective crossings. The practice of clearing and burning after the fallow period, widespread in the New Guinea Highlands, also creates favourable conditions for seed germination. Finally, it has been observed that farmers recognize the value and take care of plants issued from true seed. They name those that are selected and progressively add some to their stock of landraces (Schneider, 1995).

The highlands, a mountainous environment, exhibit a diversity of environmental, ecological, and cultural conditions (Hays, 1993). In the Baliem Valley, for example, there are two major agro-ecologies; alluvial with wetter conditions in the centre of the valley and dryer on the surrounding slopes. Farmers clearly distinguish their landraces in respect to their broad adaptation to each of these environments (Schneider, 1995). Moreover, mountainous terrain and tribal hostilities may have limited the movement of people and therefore also that of plant germplasm (Lutulele, 2001). Geographical isolation in this heterogeneous environment must have been important for generating new, locally adapted genotypes.

In lowland areas, no such precise observations have been reported, and flowering is probably a rare event in these wet, almost equatorial, conditions, although it likely does occur in drier areas or periods (P. Van Wijmeersch, pers. comm., 2011). Our results demonstrate that pairwise dissimilarities between lowland landraces are high, even higher than between different highland landraces. Recombination must also have been a major mechanism of landrace diversification in lowland agroecosystems, suggesting that here as well, seedlings appear and are selected. Alternatively, repeated introductions from tropical America, where sexuality is also a major contributor to evolutionary processes (Roullier *et al.*, 2011), could also contribute to the pattern of diversity in lowland accessions.

### ***Patterns of diversity in New Guinea: some insights into the conservation of sweet potato***

Our data showed that there is weak genetic differentiation within the New Guinea genepool. Allelic diversity in New Guinea does not appear to show strong geographical structure. Moreover, level of diversity in terms of allelic richness and rarity seems to be relatively homogeneous across geographical regions. Relatively recent (around 300-400 years ago) introduction, predominantly from one American genepool, combined with movements of clones across the island, and very active sexual recombination, are factors that may explain the near absence of geographical structure assessed with neutral markers.

These results highlight the fact that neutral nuclear markers are not the most efficient for selecting varieties to constitute core subsets. There is certainly a differentiation between lowland and



highland varieties that is based on ecological gradients. Six-hundred-seventy-eight highland varieties, representing approximately 60% of the highland collections, have already been evaluated in lowland environments. Most of them gave only very low yields or even no yield at all (Van Wijmeersch, 2001). However, neutral markers do not reflect very well patterns of variation related to environmental gradients, including divergence in adaptive traits (McKay and Latta., 2002; Hoffmann and Willi, 2008; Gebremedhin *et al.*, 2009). For New Guinea sweet potato landraces, conventional evaluation of morphological, physiological or other attributes (e.g., organoleptic traits) remains an appropriate approach for identifying useful genotypes and defining core subsets for *ex situ* conservation.

Farmers and their practices appear to be keystone actors in determining the evolutionary and adaptive dynamics of sweet potato, because of their ability to manage the crop's sexual reproductive biology. Throughout the island, subsistence farmers do not readily adopt improved varieties developed away from farming systems and local environments (Wijmeersch, 2001). It thus seems more appropriate to develop new cultivars in the different agroecological zones. The geographical distribution of allelic diversity (DAD), integrating both *ex situ* conservation and participatory breeding (Lebot *et al.*, 2005), has been proposed as a practical alternative to avoid the costs and laborious maintenance of *ex situ* collections in developing countries, and to help farmers to fashion relatively autonomous strategies for coping with ongoing global change. The DAD approach focuses on the broadening of local genetic basis, by the assemblage and evaluation in research stations of genotypes representative of the useful diversity of the species. Then, successive recombination-selection cycles on farm should be able to give rise quite rapidly to locally adapted variants. Success of such an approach depends largely on how frequently seedlings are incorporated by farmers. Generally, farmers do not really take care of seedlings for utilitarian purposes; their management of crops over many years cannot be characterized as consciously having the purpose of plant breeding. They have been taking advantage of a natural breeding/crossing process that resulted in the continual production of spontaneously germinating true sweet potato seeds. The number of seedlings produced will depend largely on features of the agroecosystem such as the length of the fallow period and the planting (or not) of varieties belonging to different self-incompatibility groups in the same field. Once produced, the extent to which plants issued from sexually produced true seed actually contribute to diversity and adaptation depends on farmers' knowledge about them and attitudes towards them. On-farm approaches to conserving the dynamics of sweet potato diversification should thus promote not only farmers' knowledge about volunteer seedlings, but also agro-ecological conditions and cultivation practices that favour their abundant production.

## **CONCLUSION**

Combining nuclear and chloroplast data, we showed that New Guinea landraces are principally derived from the Northern neotropical genepool (Caribbean and Central America), but that some South American clones may also have been introduced, either early by Polynesians themselves, or (more likely) later by whalers and missionaries and through 20<sup>th</sup>-Century movements across the Pacific. Subsequent recombination between these multiple introductions, the frequent incorporation of plants issued from true seed by farmers, and the geographical and cultural barriers constraining crop diffusion in this topographically and linguistically heterogeneous area, have generated quite rapidly an impressive number of variants, adapted to a wide range of agroecological zones through the island. Integrating such evolutionary processes in conservation strategies could provide plant scientists with a powerful tool in supporting local communities to adapt their farming systems to ongoing environmental and societal changes.

## **ACKNOWLEDGEMENTS**

This study was funded by grants from the “Pacific Funds” distributed by the French Ministry of European and Foreign Affairs. The first author, C.R., was supported by a doctoral grant from the French Ministry of Higher Education and Research. The authors are very grateful to NARI field helpers for collecting plant material. We also thank Paul Van Wijmeersch (formerly of the NARI) for useful comments on previous versions of the manuscript and for stimulating discussions. Data used in this work were produced through the technical facilities of the Centre Méditerranéen Environnement Biodiversité.

## **CONFLICT OF INTEREST**

The authors declare no conflict of interest.

## **DATA ARCHIVING**

Genotype data will be archived to Dryad (in progress).

## **SUPPLEMENTARY INFORMATION**

Supplementary information is available at Heredity’s website

## REFERENCES

- Asfaw A, Blair MW, Almekinders C (2009) Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces from the East African highlands. *Theor Appl Genet* **120**:1-12.
- Arnaud-Haond S, Duarte CM, Alberto F, Serrão EA (2007) Standardizing methods to address clonality in population studies. *Mol Ecol* **16**:5115-5139.
- Ballard C (2005) Still good to think with: sweet potato in Oceania. In: Ballard C, Brown P, Bourke RM, Harwood T (eds), *The Sweet Potato in Oceania: a Reappraisal*, Ethnology Monographs 19, Department of Anthropology, University of Pittsburgh: Pittsburgh, USA, and Oceania Monograph 56, The University of Sydney: Sydney, Australia, pp 3-13.
- Ballard C, Brown P, Bourke RM, Harwood T (2005) *The Sweet Potato in Oceania: a Reappraisal*, Ethnology Monographs 19, Department of Anthropology, University of Pittsburgh, Pittsburgh and Oceania Monograph 56, The University of Sydney, Sydney.
- Barrau J (1957) L'énigme de la patate douce en Océanie. *Etudes d'Outre-Mer* **40** : 83-87.
- Bayliss-Smith T, Golson J, Hughes P, Blong R, Ambrose W (2005) Archaeological evidence for the Ipomoea Revolution at Kuk swamp, upper Waghi Valley, Papua New Guinea. In: Ballard C, Brown P, Bourke RM, Harwood T (eds) *The Sweet Potato in Oceania: a Reappraisal*, Ethnology Monographs 19, Department of Anthropology, University of Pittsburgh, Pittsburgh, USA and Oceania Monograph 56, The University of Sydney: Sydney, Australia, pp 3-13.
- Bourke RM (2005) The continuing Ipomoea Revolution in Papua New Guinea. In: Ballard C, Brown P, Bourke RM, Harwood T (eds) *The Sweet Potato in Oceania: a Reappraisal*, Ethnology Monographs 19, Department of Anthropology, University of Pittsburgh: Pittsburgh, USA and Oceania Monograph 56, The University of Sydney: Sydney, Australia, pp 171-180.
- Bourke M (2009) Sweet potato in Oceania. In: Loebeinstein G, Thottappilly G (eds) *The Sweet Potato*, Springer-Verlag.
- Bourke RM, Harwood T (2009) *Food and Agriculture in Papua New Guinea*. ANU E Press, The Australian National University: Canberra.
- Brown AHD (1989) Core collections - a practical approach to genetic resources management. *Genome* **31**:818-824.
- Bulmer R (1965) Beliefs concerning the propagation of new varieties of sweet potato in two New Guinea Highland societies. *J Polynesian Soc* **74**:237-239.
- Burle ML, Fonseca JR, Kami JA, Gepts P (2010) Microsatellite diversity and genetic structure among common bean (*Phaseolus vulgaris* L.) landraces in Brazil, a secondary centre of diversity. *Theor Appl Genet* **121**:801-813.
- Caillon S, Quero-Garcia J, Lescure JP, Lebot V (2006) Nature of taro (*Colocasia esculenta* (L.) Schott) genetic diversity prevalent in a Pacific Ocean island, Vanua Lava, Vanuatu. *Genet Resour Crop Evol* **53**:1273-1289.
- Clarke AC, Burtenshaw MK, McLenachan PA *et al.*, (2006) Reconstructing the origins and dispersal of the Polynesian bottle gourd (*Lagenaria siceraria*). *Mol Biol Evol* **23**:893-900.
- Clark LV, Jasieniuk M (2011) polysat: an R package for polyploid microsatellite analysis. *Mol Ecol Resour* **11**:562-566.
- Delêtre M (2010) The ins and outs of cassava diversity in Gabon, Central Africa: a pluridisciplinary approach to the dynamics of genetic diversity of *Manihot esculenta* Crantz (Euphorbiaceae). PhD Dissertation, Trinity College, University of Dublin (<http://tel.archives-ouvertes.fr/tel-00623219/fr/>).
- Elias M, Penet L, Vindry P, McKey D, Panaud O, Robert T (2001) Unmanaged sexual reproduction and the dynamics of genetic diversity of a vegetatively propagated crop plant, cassava (*Manihot esculenta* Crantz), in a traditional farming system. *Mol Ecol* **10**:1895-1907.
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol Bioinformatics* **1**:47-50.
- Fajardo DD, La Bonte DR, Jarret RL (2002) Identifying and selecting genetic diversity in Papua New Guinea sweet potato *Ipomoea batatas* (L.) Lam. germplasm collected as botanical seed. *Genet Resour Crop Evol* **49**:463-470.
- Gichuki ST, Berenyi M, Zhang D *et al.* (2003) Genetic diversity in sweetpotato [*Ipomoea batatas* (L.) Lam.] in relationship to geographic sources as assessed with RAPD markers. *Genet Resour Crop Evol* **50**:429-437.
- Golson J (1977) No room at the top: agricultural intensification in the New Guinea Highlands. In: Allen J, Golson J, Jones R (eds) *Sunda and Sahul*, Academic Press: London, pp 601-638.
- Gorecki PP (1986) Human occupation and agricultural development in the Papua New Guinea Highlands. *Mountain Research and Development* **6**:159-166.

- Green RC (2005) Sweet potato transfers in Polynesian prehistory. In Ballard C, Brown P, Bourke RM, Harwood T (eds) *The Sweet Potato in Oceania: a Reappraisal*, Ethnology Monographs 19, Department of Anthropology, University of Pittsburgh: Pittsburgh, USA and Oceania Monograph 56, The University of Sydney: Sydney, Australia, pp 43-62..
- Gebremedhin B, Ficitola GF, Naderi S, Rezaei HR, Maudet C, Rioux D, et al (2009) Frontiers in identifying conservation units: from neutral markers to adaptive genetic variation. *Anim Conserv* **12**:107-109.
- Gunn BF, Baudouin L, Olsen KM (2011) Independent origins of cultivated coconut (*Cocos nucifera* L.) in the Old World tropics. *PLoS ONE* **6** (6). e21143.doi:10.1371/journal.pone.0021143..
- Haberle SG (1998) Dating the evidence for prehistoric agricultural change in the highlands of New Guinea: the last 2000 years. *Archaeology in Oceania* **47**:1-19.
- Haberle SG, Atkin G (2005) Needles in a haystack: searching for sweet potato (*Ipomoea batatas* (L.) Lam.) in the fossil pollen record. In: Ballard C, Brown P, Bourke RM, Harwood T (eds) *The Sweet Potato in Oceania: a reappraisal*, Ethnology Monographs 19, Department of Anthropology, University of Pittsburgh: Pittsburgh, USA and Oceania Monograph 56, The University of Sydney: Sydney, Australia, pp 25-34.
- Hays TE (1993) "The New Guinea Highlands": Region, culture area, or fuzzy set? *Curr Anthropol* **34**:141-64.
- Hoffmann AA, Willi Y (2008) Detecting genetic responses to environmental change. *Nat Rev Genet* **9**:421-432.
- Hinkle AE (2007) Population structure of Pacific *Cordyline fruticosa* (Laxmanniaceae) with implications for human settlement of Polynesia. *Am J Bot* **94**:828-839.
- Jombart T (2008) Adegnet: R package for the multivariate analysis of genetic markers. *Bioinf* **24**: 1403-1405.
- Jombart T, Devillard S, Balloux F (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* **11**:94.
- Kreike CM, Van Eck HJ, Lebot V (2004) Genetic diversity of taro, *Colocasia esculenta* (L.) Schott, in Southeast Asia and the Pacific. *Theor Appl Genet* **109**:761-768.
- Ladefoged TN, Graves MW, Coil JH (2005) The introduction of sweet potato in Polynesia: Early remains in Hawai'i. *J Polyn Soc* **114**:359-373.
- Lebot V (1992) Genetic vulnerability of Oceania's traditional crops. *Exp Agric* **28**:309-323.
- Lebot V (1999) Biomolecular evidence for plant domestication in Sahul. *Genet Resour Crop Evol* **46**:619-628.
- Lebot V, Ivancic A, Abraham K (2005) The geographical distribution of allelic diversity, a practical means of preserving and using minor root crop genetic resources. *Exp Agric* **41**:475-489.
- Lutulele RP (2001) Sweet potato variety developments in the PNG highlands: implications for future research and extension focus. In: Bourke RM, Allen MG, Salisbury JD (eds) *Food Security for Papua New Guinea*, Proceedings of the Papua New Guinea Food and Nutrition 2000 Conference (ACIAR Proceedings 99), University of Technology: Lae, pp 683-688.
- McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends Ecol Evol* **17**:285-291.
- McKey D, Elias M, Pujol B, Duputie A (2010) The evolutionary ecology of clonally propagated domesticated plants. *New Phytol* **186**:318-332.
- Nakanishi T, Kobayashi M (1979) Geographic distribution of cross incompatibility groups in sweet potato. *Incompatibility Newsl* **11**:72-75.
- Ollitrault P, Luro F (2001) Citrus. In: Charrier A, Jacquot M, Hamon S, Nicolas D (eds) *Tropical Plant Breeding*, CIRAD: Paris, pp 55-77.
- Perrier X, De Langhe E, Donohue M, Lentfer C, Vrydaghs L, Bakry F, et al. (2011) Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proc Natl Acad Sci USA* **108**:11311-11318.
- Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conserv Biol* **12**:844-855.
- Pickersgill B (1998) Crop introductions and the development of secondary areas of diversity. In: Prendergast HDV, Etkin NL, Harris DR, Houghton PJ (eds), *Plants for Food and Medicine*, Royal Botanic Gardens: Kew, pp 93-105.
- R Development Core Team (2009) R: a language and environment for statistical computing, reference index version 2.8.1. Vienna, Austria: R Foundation for Statistical Computing.
- Romão RL (2000) Northeast Brazil: A secondary centre of diversity for watermelon (*Citrullus lanatus*). *Genet Resour Crop Evol* **47**:207-213.
- Roullier C, Rossel G, Tay D, McKey D, Lebot V (2011) Combining chloroplast and nuclear microsatellites to investigate origin and dispersal of New World sweet potato landraces. *Mol Ecol* **20**:3963-3977.
- Sardos J, McKey D, Duval MF, Malapa R, Noyer JL, Lebot V (2008) Evolution of cassava (*Manihot esculenta* Crantz) after recent introduction into a South Pacific Island system: the contribution of sex to the diversification of a clonally propagated crop. *Genome* **51**:912-921.

- Sardos J, Noyer JL, Malapa R, Bouchet S, Lebot V (2011) Genetic diversity of taro (*Colocasia esculenta* (L.) Schott) in Vanuatu, Oceania: an appraisal of the distribution of allelic diversity (DAD) with SSR markers. *Genet Resour Crop Evol* **59**:805-820.
- Scaglione R, Soto K (1994) A prehistoric introduction of sweet potato in New Guinea? In Strathern AJ, Sturzenhofecker G (eds) *Migration and Transformation: Regional Perspectives on New Guinea*. ASAO Monograph 15, University of Pittsburgh Press: Pittsburgh, pp 257-294.
- Scarcelli N, Tostain S, Vigouroux Y, Agbangla C, Dai-nou O, Pham JL (2006) Farmers' use of wild relative and sexual reproduction in a vegetatively propagated crop. The case of yam in Benin. *Mol Ecol* **15**:2421-2431.
- Schneider J (1995) Farmer practices and sweet potato diversity in highland New Guinea. In: Schneider J (eds) *Indigenous Knowledge in Conservation of Crop Genetic Resources: Proceedings of an International Workshop held in Cisarua, Bogor, Indonesia, CIP-ESEAP/CRIFC*: Bogor, pp 63-70.
- Van Wijmeersch P (2001) The status of sweet potato variety evaluation in PNG and recommendations for further research. In: Bourke RM, Allen MG, Salisbury JD (eds) *Food Security for Papua New Guinea*, Proceedings of the Papua New Guinea Food and Nutrition 2000 Conference (ACIAR Proceedings 99), University of Technology: Lae, pp 683-688.
- Wiessner P (2005) Social, symbolic, and ritual roles of the sweet potato in Enga, from its introduction until first contact. In: Ballard C, Brown P, Bourke RM, Harwood T (eds) *The Sweet Potato in Oceania: a Reappraisal*, Ethnology Monographs 19, Department of Anthropology, University of Pittsburgh: Pittsburgh, USA, and Oceania Monograph 56, The University of Sydney: Sydney, Australia, pp 121-130.
- Yen DE (1974) *The Sweet Potato in Oceania: An Essay in Ethnobotany*. Bernice P. Bishop Museum 236, Bishop Museum Press, Honolulu: Hawaii.
- Zhang DP, Ghislain M, Huamán Z, Golmirzaie A and Hijmans R (1998) RAPD variation in sweet potato (*Ipomoea batatas* (L.) Lam.) cultivars from South America and Papua New Guinea. *Genet Resour Crop Evol* **45**:271-277.
- Zhang DP, Rossel G, Kriegner A, Hijmans R (2004) AFLP assessment of diversity in sweetpotato from Latin America and the Pacific region: Its implications on the dispersal of the crop. *Genet Resour Crop Evol* **51**:115-120.

## Supplementary material 1: Assigning to admixed original populations with DAPC and Bayesian clustering analysis

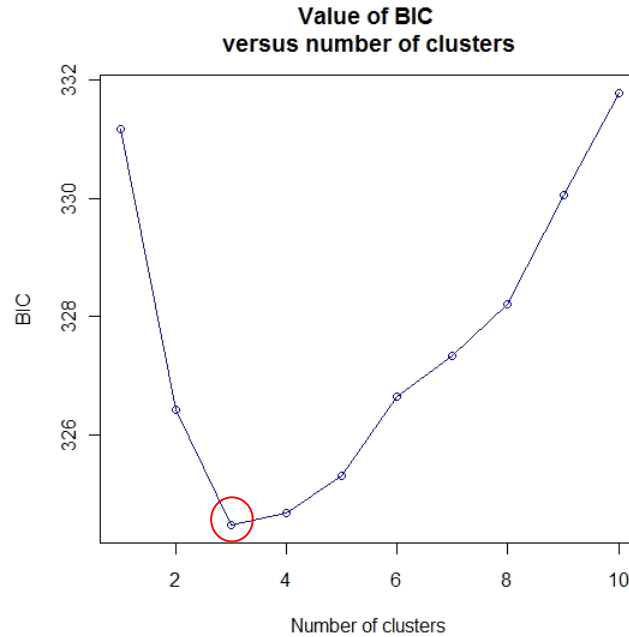
Classical genetic methods of assignment are based on the multilocus genotype of an individual and the expected probabilities of that genotype occurring in each of the potential sources. These “model-based” methods mostly rely on restrictive explicit assumptions (populations at Hardy-Weinberg equilibrium and linkage equilibrium) and were essentially developed for data on diploid (and haploid) genotypes (Manel *et al.*, 2005). Polyploid microsatellites dataset may be scored as binary data (presence-absence), but it remains unclear how these model-based methods are relevant for treating such kinds of data.

The software STRUCTURE (Prichard *et al.*, 1999), a Bayesian model-based clustering algorithm, has been recently modified to handle polyploid data and allele copy number ambiguity, under the assumption of full autopolyploid inheritance (Falush *et al.*, 2007). Commonly used to identify genetic clusters in a dataset, STRUCTURE may also be used to assign additional individuals of unknown origin to their source populations. By pre-specifying source populations, the algorithm estimates ancestry for additional individuals, updating allele frequencies using only those from source populations.

The Discriminant Analysis of Principal Components (DAPC), a non model-based method recently developed and implemented in the adegenet R packages (Jombart, 2008), provides an efficient description of genetic clusters using a few synthetic variables (called the discriminant functions). This multivariate analysis seeks linear combinations of the original variables (alleles) which show differences between groups as best as possible while minimizing variation within clusters. Contrary to traditional methods such as PCA or PCoA, which focus on the entire genetic variation, DAPC yields linear combinations of the original variables (alleles) which maximize differences between groups while minimizing variation within clusters. Based on the retained discriminant functions, the analysis derives probabilities for each individual of membership in each of the different groups. This coefficient can be interpreted as “genetic proximity” of individuals to the different clusters. It is possible to construct the linear model and obtain synthetic variables on a given dataset (the source populations), then add supplementary individuals (that were not used in constructing the model) and derive for each one a membership probability to original source populations. These coefficients might provide an “assignment measure” of individuals to predefined groups, comparable with ancestry value derived by the STRUCTURE analysis.

In this study, both methods were performed to investigate the potential source(s) of New Guinean landraces in tropical America. Therefore, we first ran these methods with the tropical America dataset, excluding New Guinean samples, which were added as supplementary individuals. However, both methods require prior groups to be achieved.

**K-means clustering and the DAPC method:** The adegenet package allows running the sequential K-means clustering algorithm, and comparing the different clustering solutions using the Bayesian Information Criterion (BIC) (after transforming the data using a PCA, notably to reduce the numbers of variables and speed the clustering algorithm) to identify an optimal number of genetic clusters to describe the data. We ran the K-means clustering algorithm for  $K = 2$  to  $K = 10$  on the tropical America dataset. Based on this analysis, three genetic clusters were considered optimal to describe the data.

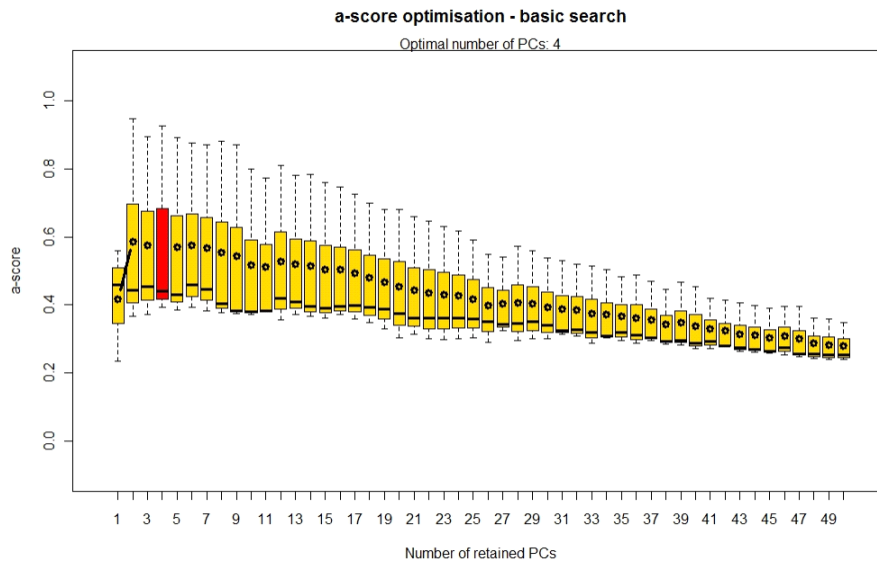


**Figure 1:** Inference of the number of clusters in the DAPC performed on the tropical America dataset. A K value of 3 (the lowest BIC value) represents the best summary of the data.

The three genetic clusters are geographically restricted: Two clusters (K2 and K3) group mostly individuals from the Northern region (90.6 %) while cluster K1 groups mostly those from the Southern region (81.4 %). The grouping obtained for  $K = 2$  also provides an accurate description of the tropical America dataset: the cluster *K1* mainly contained samples from the Southern genepool (79%), and cluster *K2* mostly those from the Northern genepool (91%).

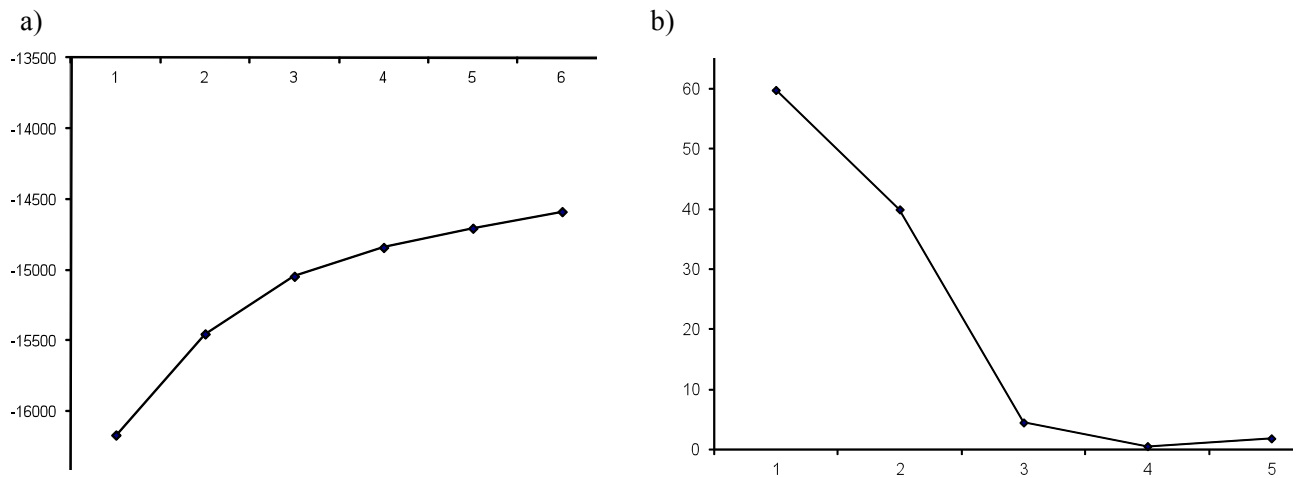
DAPC relies on data transformation using principal component analysis (PCA) as a prior step. Retaining too many PCs can lead to overfitting the discriminant functions, which could model any structure and virtually discriminate any set of clusters. Adegenet proposes an optimization procedure to evaluate the optimal numbers of PCs to retain. The procedure is based on the calculation of the  $\alpha$ -score, which measures the difference between the proportion of successful reassignment of the analysis (observed discrimination) and values obtained using random groups (random discrimination). The number of retained PCs can be chosen so as to optimize the  $\alpha$ -score.

The optimization  $\alpha$ -score graph (Figure 2) shows that only few PCs need to be retained for the assignment analysis. We tested DAPC for both groupings ( $K = 2$  and  $K = 3$ ), retaining between 4 and 10 PCs, for prior data transformation. Assignment results were globally congruent and we present in the article results obtained with 5 PCs retained (29.3 of the total variance).



**Figure 2:** Optimization  $\alpha$ -score graph

**Bayesian clustering method:** We also used a Bayesian clustering method on the tropical America dataset to predefine groups in tropical America that may be used for the assignment of New Guinea accessions. We ran STRUCTURE for  $K = 2$  to  $K = 6$ , using the admixture model, correlated allele frequencies, 50 000 burn-in iterations and 150 000 Markov chain-Monte Carlo steps and data coding for handling genotype ambiguity for co-dominant markers in polyploids (Falush *et al.*, 2007). We then plotted the  $\Delta K$  Evanno criterion (Evanno *et al.*, 2005) to identify the optimal number of clusters to describe the data.



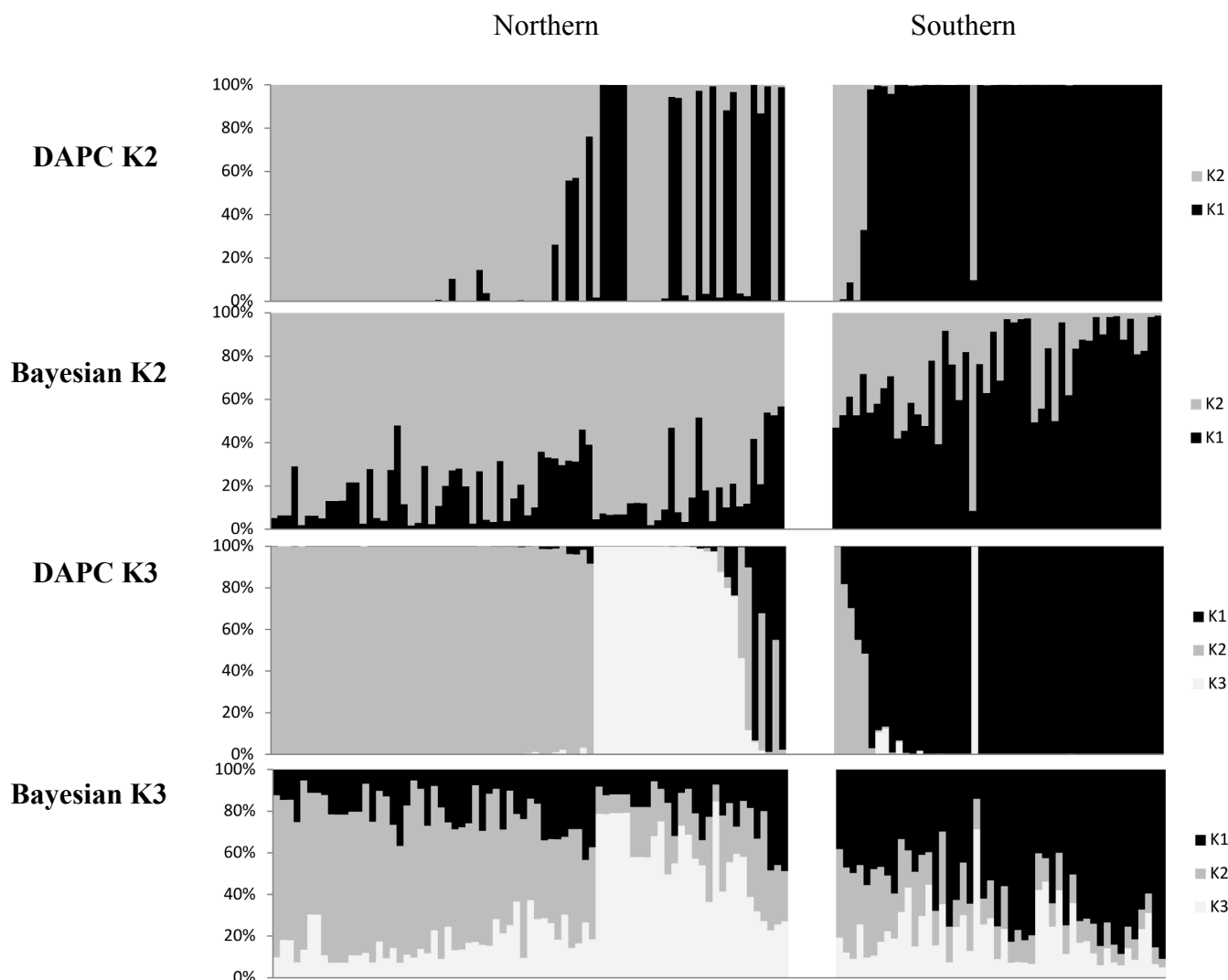
**Figure 3:** a) Variation of the posterior log-probability of the data as a function of the number of clusters  $K$ . b) Variation of  $\Delta K$  values

Following this method, the optimal number of clusters to describe the data was unclear. We then retained groupings obtained for  $K = 2$  and  $K = 3$  (as determined by the BIC criterion) and compared their composition with those obtained with the DAPC method.

### Results:

**Figure 3:** Tropical America landraces membership probabilities (DAPC analysis) or ancestry value (Bayesian analysis) to K1 and K2 (DAPC K2 or Bayesian K2) or K1, K2 and K3 (DAPC K3 or Bayesian K3) clusters. Each individual is represented as a vertical bar, with colours corresponding to probabilities of membership in K1 (black), K2 (dark gray), and K3 (light gray).





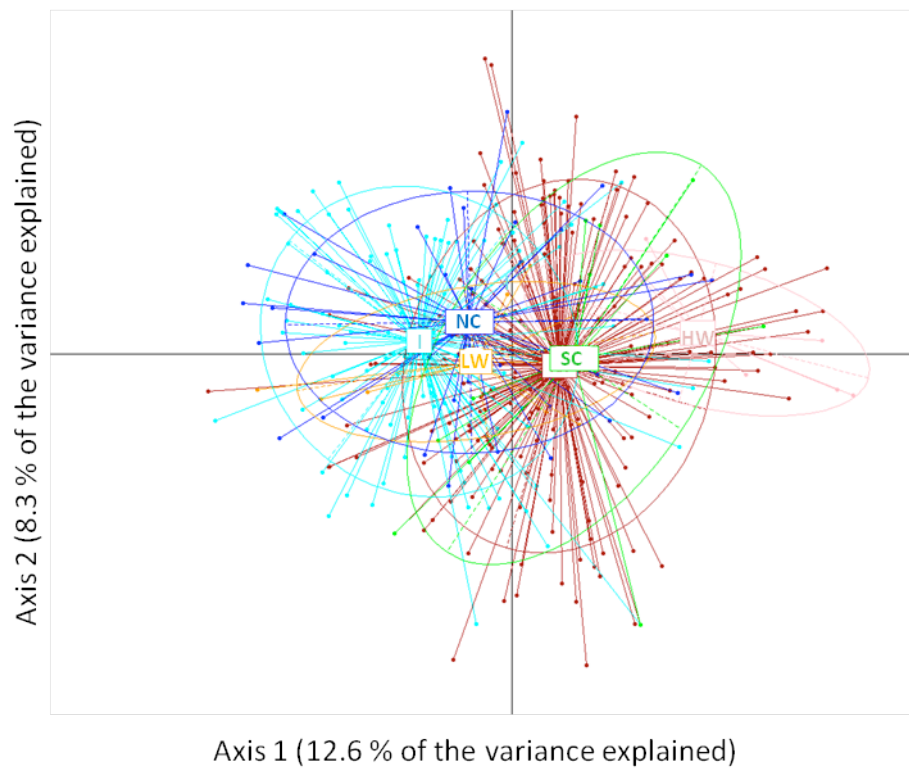
Both methods gave globally congruent results: Neotropical sweet potatoes are characterized by two distinct genetic groups, geographically circumscribed: one genetic group corresponds to most of the accessions from the Southern region and the other one to those from the Northern region. For  $K = 3$ , a sub-structure is revealed in which Northern region accessions are split into two genetic groups.

However, the phylogeographic pattern is not so clear-cut: indeed, in the Southern region, we detected some accessions clearly attributed to the nuclear cluster(s) characteristic(s) of the Northern region (K2, or K2 and K3) or with a mixed genetic constitution. Also, in the Northern region, we identified several accessions attributed to cluster K1 or with a mixed composition. As we already discussed in a previous paper (Roullier *et al.*, 2011), this situation suggests that Neotropical sweet potatoes are characterized by two original differentiated genepools (probably related to independent domestications in each region) and that clones were secondarily exchanged between both regions and then recombined with local material. This scenario is also supported by chloroplast data (Roullier *et al.*, 2011). This situation of admixture is well underlined by the Bayesian clustering results, for which most of the neotropical accessions exhibited a mixed genetic constitution. With the DAPC method, individual assignment was more “contrasted”, only few individuals exhibiting a mixed constitution.

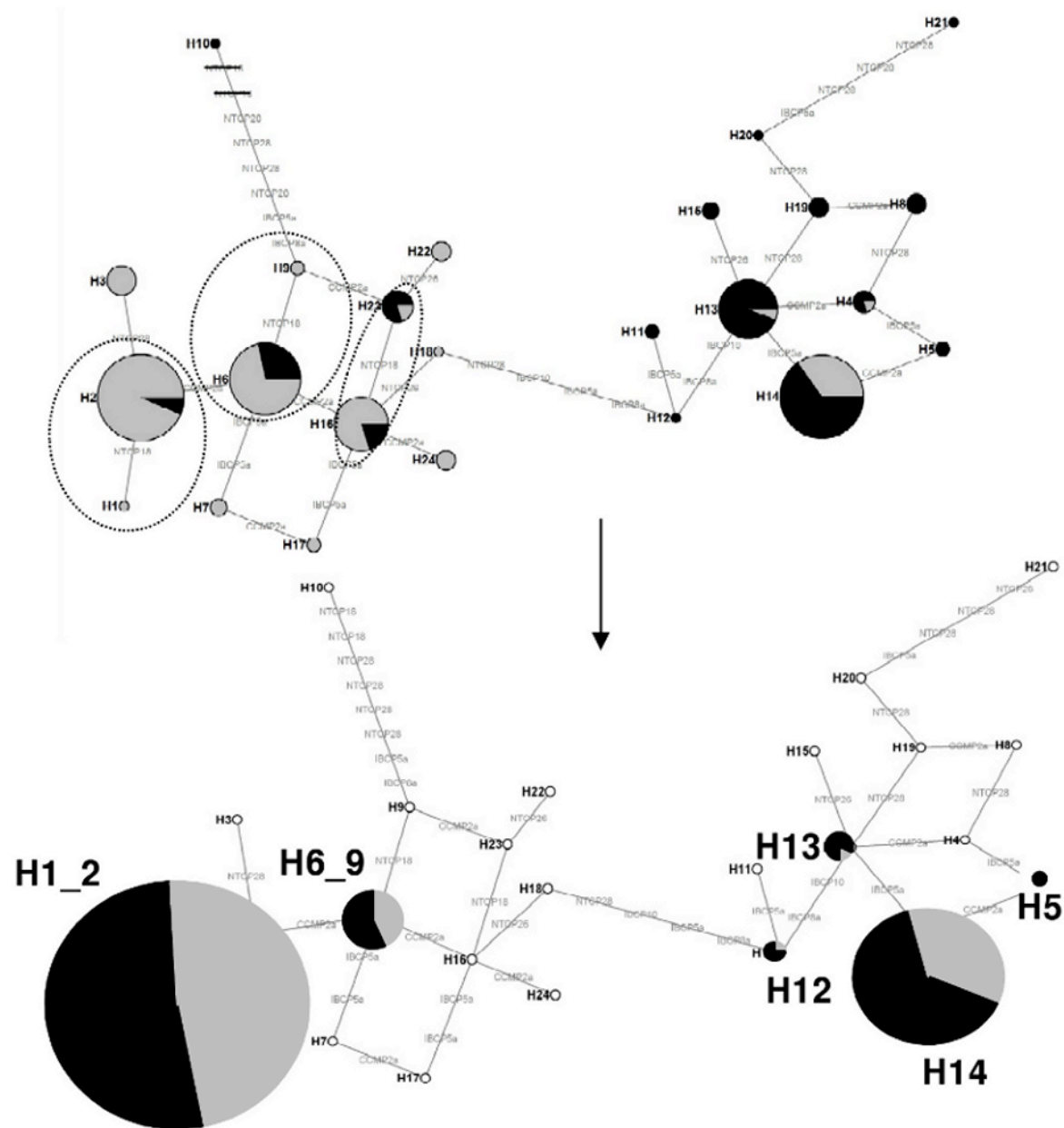
Thus, it was difficult to use Bayesian clustering results on the tropical America dataset to predefine groups for assignment of the New Guinean accessions. We preferred to use the genetic grouping (and not the *a priori* regional grouping) inferred by the K-means clustering for  $K = 2$ , which provides an accurate and simple summary of the neotropical dataset, to perform assignment of New Guinean accessions.

- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**:2611-2620.
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164**:1567–1587.
- Jombart T (2008) Adegenet: R package for the multivariate analysis of genetic markers. *Bioinf* **24**: 1403-1405.
- Manel S, Gaggiotti OE, Waples RS (2005) Assignments methods: matching biological questions with appropriate techniques. *Trends Ecol Evol* **20**:136-142.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**:945–959.
- Roullier C, Rossel G, Tay D, McKey D, Lebot V (2011) Combining chloroplast and nuclear microsatellites to investigate origin and dispersal of New World sweet potato landraces. *Mol Ecol* **20**:3963-3977.

**Supplementary material 2:** Principal Coordinate Analysis based on the Lynch distance between genotypes. Individuals are labelled following pre-defined geographical areas: NC for the lowlands of the northern coast of Papua New Guinea in blue, I for the islands region of Papua New Guinea in magenta, HP for the highlands of Papua New Guinea in red, HW for the highlands of West New Guinea in pink, SC for the lowlands of the southern coast of Papua New Guinea in green, and LW for the lowlands of West New Guinea in orange.



**Supplementary material 3:** Genetic relationships and distribution of chloroplast haplotypes across tropical America and New Guinea: comparison of the median-joining network obtained for tropical American accessions using seven chloroplast SSRs (data already published in Roullier *et al.* [2011]) and the network adapted for New Guinean accessions using only six polymorphic chloroplast SSRs. H1 and H2 became one haplotype (H1\_2), as did H6 and H9 (H6\_9) and H23 and H16 (H23\_16). One-bp mutation events are indicated by the name of the locus concerned. Circle size is proportional to the number of accessions that carry a given haplotype. For each haplotype, the proportion of accessions from Southern and Northern neotropical genepools that carry the haplotype is represented in black and grey, respectively. Similarly, the proportion of highland and lowland New Guinea accessions carrying the haplotype are indicated in black and grey, respectively.



## *Chapitre 6*

*Article : Les collections historiques révèlent des patrons de diffusion de la patate douce (Ipomoea batatas (L.) Lam.) en Océanie brouillé par les mouvements et la recombinaison des plantes modernes*

*Historical collections reveal patterns of diffusion of sweet potato in Oceania obscured by modern plant movements and recombination*

**1. Title:** Historical collections reveal patterns of diffusion of sweet potato in Oceania obscured by modern plant movements and recombination

**2. Authors' names:** C. Roullier<sup>1,2</sup>, L. Benoit<sup>2</sup>, D. McKey<sup>2,3</sup> and V. Lebot<sup>1</sup>

**3. Full postal address:**

<sup>1</sup> CIRAD-BIOS, AGAP, Avenue Agropolis, 34398 Montpellier Cedex 5, France

<sup>2</sup> CEFE, UMR 5175 CNRS, 1919 route de Mende, 34293 Montpellier Cedex 5, France

<sup>3</sup> Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier cedex 5, France; Institut Universitaire de France

**4. Corresponding author:**

Roullier Caroline, CIRAD, Avenue Agropolis, 34398 Montpellier Cedex 5 France. CEFE, 1919 route de Mende, 34293 Montpellier Cedex 5, France; E-mail address: [carolineroullier@hotmail.fr](mailto:carolineroullier@hotmail.fr)

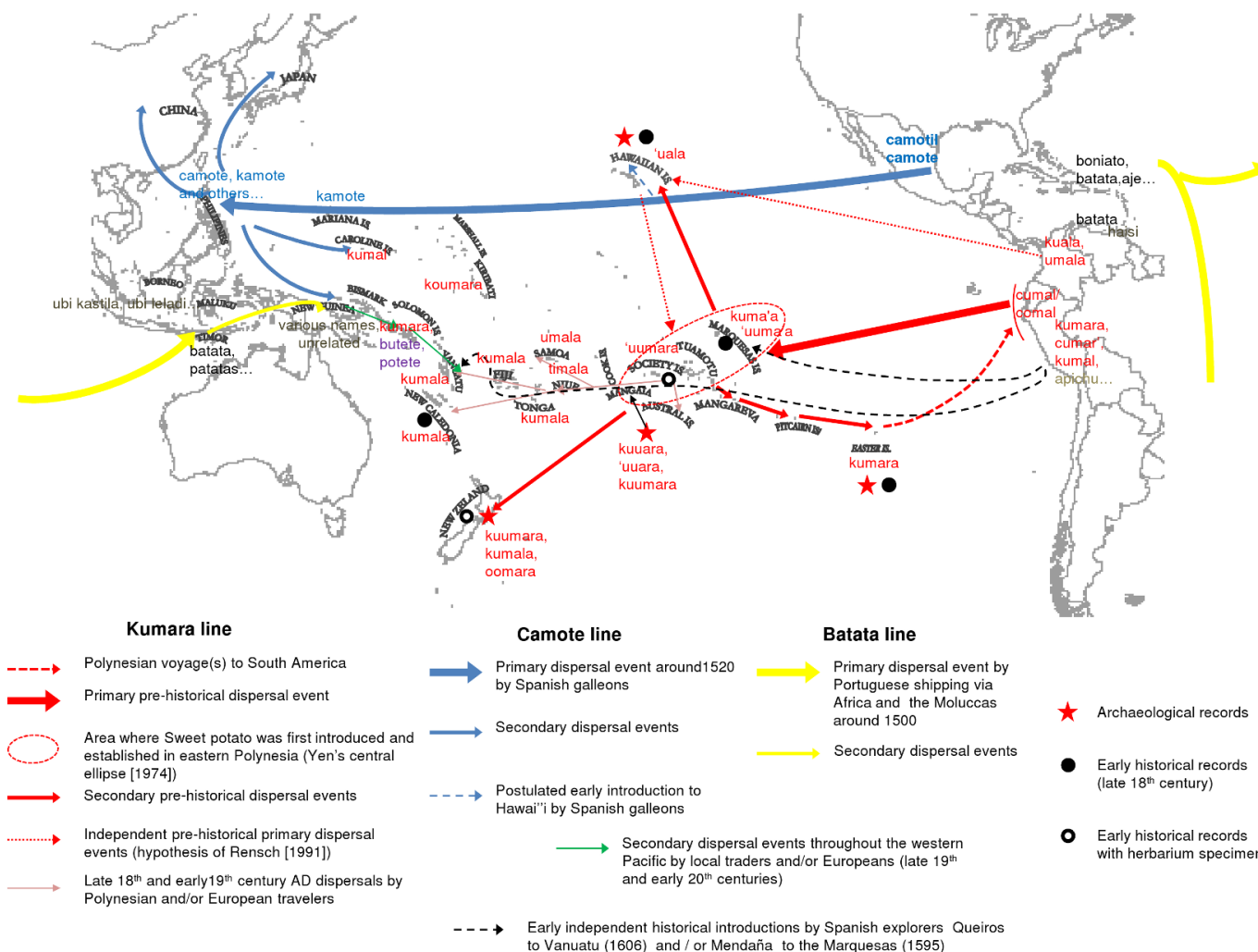
**5. A condensed version of this chapter was submitted to PNAS. Accepted with minor revisions.**

## ABSTRACT

The history of sweet potato in the Pacific has long been an enigma. Archaeological, linguistic and ethnobotanical data suggest that prehistoric human-mediated dispersal events contributed to the distribution in Oceania of this American domesticate. According to the “tripartite hypothesis”, sweet potato was introduced into Oceania from South America in pre-Columbian times, and was then later newly introduced, and diffused widely across the Pacific, by Europeans via two historically documented routes from Mexico and the Caribbean. Although sweet potato is the most convincing example of putative pre-Columbian connections between human occupants of Polynesia and South America, the search for genetic evidence of pre-Columbian dispersal of sweet potato into Oceania has been inconclusive. Our study attempts to fill this gap. Using complementary sets of markers (chloroplast and nuclear microsatellites), and both modern and herbarium samples, we test the tripartite hypothesis. Our results provide strong support for prehistoric transfer(s) of sweet potato from South America (Peru-Ecuador region) into Polynesia. They also document a temporal shift in the pattern of distribution of genetic variation in sweet potato in Oceania. Later re-introductions, accompanied by recombination between distinct sweet potato genepools, have reshuffled the crop's initial genetic base, obscuring primary patterns of diffusion and at the same time giving rise to an impressive number of local variants. Moreover, our study shows that phenotypes, names and neutral genes do not necessarily share completely parallel evolutionary histories. Multidisciplinary approaches thus appear necessary for accurate reconstruction of the intertwined histories of plants and humans.

**Figure 6-1: The pre-historic and historical dispersal of sweet potato in Oceania as postulated by the tripartite hypothesis.**

This map summarizes the tripartite hypothesis as updated by Green (2005). Dispersal events and the terms commonly used to designate sweet potato in the different regions (compiled mostly from Yen 1974) are represented. Archaeological and early historical records providing strong evidence for the presence of sweet potato in Oceania in the pre-historical period are also indicated.



## INTRODUCTION

Recently, the reconstruction of human mobility and cultural evolution in the Pacific has focused not only on the dispersal of humans themselves (1–3), but also examined that of particular species closely associated with human movements (4–12). Studies of commensal animals (*Rattus exulans*) (4, 5), domesticated plants (bottle gourd, *Lagenaria siceraria* (6, 7); banana, *Musa* spp.) (8)) and animals (chicken, *Gallus gallus* (9, 10); pig, *Sus scrofa* (11)) and pathogens (*Helicobacter pylori*) (12) have been successfully employed to reconstruct ancient patterns of human migrations and interactions.

There has long been debate regarding the question of connection(s) between Polynesia and the Americas in prehistoric times (13, 14). Sweet potato (*Ipomoea batatas* [L.] Lam.), a tropical root crop domesticated in tropical America (15–17), is undoubtedly the model most convincingly suggestive of such contacts according to archaeological, linguistic, and historical evidences (18, 19, 20) (Figure 6-1). The oldest direct evidence of the presence of sweet potato in Oceania is provided by the carbonized sweet potato tuber remains from the Tangatatan rock-shelter on the island of Mangaia, dated by context at around 1000–1100 A.D. (21). Further archaeological evidence of the pre-historical presence of sweet potato has been found in Hawaii (charcoal fragments dated around 1290–1430 A.D. (22) and sweet potato microfossils (23)), Easter Island (tubers dated by context around 1400–1600 A.D., and archaeological analysis of landscapes and field systems (24)) and New Zealand (sweet potato microfossils (25)). These data suggest that sweet potato was widely distributed in Polynesia well before European contact. European travellers' early historical accounts reinforce these conclusions. Indeed, sweet potato was recorded by Roggenven in Easter Island in 1722, by J. Banks and D. Solander (botanists of James Cook's first voyage in the Pacific) in Tahiti and in New Zealand in 1769, by Cook in Hawaii in 1778, and by E. Marchand in the Marquesas in 1790 (16, 17). Although natural long-distance dispersal by seeds has been sometimes considered (26, 27), it is likely that humans introduced sweet potato from South America to Polynesia, as suggested by the lexical similarity between terms for sweet potato in Polynesian languages ("kuumala" and its derivatives), and the terms for this plant ("kumara", "cumar" or "cumal") found among Quechua speakers of north-western South America (19, 20, 28). In a comprehensive review tying archaeological and linguistic data to information on winds and currents, Green (19) proposed that Polynesian voyagers probably left Easter Island around 1000 A. D. and sailed to South America, reaching the gulf of Guayaquil (Peru/Ecuador area), where they obtained the sweet potato. Sailing back to Polynesia, they may have made landfall somewhere between Mangaia Island and the Marquesas in the central ellipse (18). The plant then rapidly diffused throughout Polynesia, between 1000 and 1150 A.D., in already populated islands such as Hawaii, Easter Island, and some other islands of the central ellipse, and in New Zealand, around 1150–1250 A.D. with the original colonists (19).



Other independent prehistoric introductions (from northern Colombia and even Central America) have also been hypothesized but these lack support (27, 29). The possibility of an early westward dispersal of sweet potato carried by Polynesians to Tonga, Samoa and eastern Melanesia is suggested by early historical accounts (30). These prehistoric transfers of sweet potato in Oceania constitute the Kumara line of the tripartite hypothesis (Figure 6-1).

Sweet potato was also introduced into the western Pacific early in historical times (18, 19). Spanish “Manila-Acapulco” galleons introduced the Mesoamerican Camote to the Philippines around 1500 (Camote line), while Portuguese traders introduced to present-day eastern Indonesia the Batata from the Caribbean and Central America (Batata line). From these points, secondary dispersal events, mediated by local traders, European travellers or both, may have distributed sweet potato in the western Pacific, probably early in the New Guinea highlands (around 1700), and much later (in the mid-19<sup>th</sup> century) into western and eastern Melanesia (Figure 6-1). Two additional European introductions may also have contributed to early introductions into the Pacific, one mediated by Mendaña to the Marquesas and Solomon Islands in 1595 and another by Queiros in Espiritu Santo (part of present-day Vanuatu) in 1606 (31). Thus, ethnobotanists, linguists, anthropologists and archaeologists have provided strong support for the “tripartite hypothesis” which posits that both prehistoric (the Kumara line from South America) and historic transfers (the Camote line from Central America and the Batata line from the Caribbean) are required to explain the sweet potato distribution in Oceania.

Regrettably, previous attempts to elucidate the origins of sweet potato in Oceania using molecular markers have not yet adequately tested this hypothesis (32-35). Zhang et al. (33) showed that varieties from the Pacific were more similar genetically to Mexican accessions than to those from Peru-Ecuador, supporting a Mesoamerican origin rather than early Polynesian transfers from South America. However, these studies suffered from a very restricted sample, including very few Polynesian varieties and used markers of limited phylogeographic resolution. More recently, Clarke et al. (35) analyzed with AFLP markers around 300 Oceanian varieties including 40 Polynesian varieties. However, this study did not use a consistent sampling of the tropical America source gene pools and thus failed to firmly retrace the origin of extant Pacific lineages. An additional limitation to retrace sweet potato history in Oceania comes from that contemporary patterns of genetic variation of sweet potato in Oceania probably do not reflect initial patterns. From all putative entry points, 20<sup>th</sup>-century intensification of human movements and exchanges across the Pacific probably extended the range of sweet potato in Oceania, as well as re-distributing it, potentially blurring initial introduction patterns (35). Moreover, primary introduction did not always lead to widespread adoption and further secondary diversification: the importance of particular crops changes over space and time (16). Sweet potato was probably quite rapidly adopted and became predominant in “marginal” areas, where its tolerance to a broad range of edaphic and climatic conditions conferred advantages over the other two traditional root crops of the Pacific: taro (*Colocasia esculenta* [L.] Schott) and the greater yam (*Dioscorea alata* L.). Sweet potato probably encountered success soon after its introduction into the New Guinea highlands, Hawaii, New Zealand and Easter Island. In lowland humid tropical habitats, this crop probably long remained of only secondary importance (36). Societal changes (urbanization, loss of tradition and expanding population, among others) that began in the second half of the 20<sup>th</sup> century are still modifying the distribution of the species (36). Consequently, genetic signatures of initial introductions can be at least partially obscured by later introductions of conspecifics from a different genetic background (10, 37). It is thus necessary to compare geographic patterns in genetic variation through time, by analyzing both modern and ancient samples. Archaeological remains of sweet potato are scarce in the Pacific and are highly unlikely to contain preserved DNA. However, herbarium specimens, collected by early European explorers and naturalists and later by botanists and agronomists, are invaluable sources of material for tracing early movements of gene pools, complementing sampling of modern accessions, more affected by the intensification of germplasm exchanges during the 20<sup>th</sup> century.

This study attempts to fill the gap of genetic data in the reconstruction of sweet potato history in Oceania. Using complementary sets of markers (chloroplast and nuclear microsatellites) and a broad sampling of both modern and herbarium accessions, we tested whether genetic evidence supports or falsifies the tripartite hypothesis proposed to explain the distribution of sweet potato in Oceania.

## MATERIALS AND METHODS

### *Sampling*

We analyzed a total of 1245 sweet potato accessions originating from tropical America and Oceania and including some accessions from South-East Asia and Madagascar. We recovered 1188 “modern” accessions from several important *ex situ* collections assembled between 1960 and 2011:

- The Yen Sweet Potato collection (Yen) assembled in the 1960s by D. Yen and conserved in the National Institute for Crop Sciences in Tsukuba, Japan.
- The Highlands and Lowlands collections (HAES, and LAES) of the National Agricultural Research Institute (NARI), Papua New Guinea, collected mostly in the 1980s, and maintained at the Highlands Agricultural Experiment Station (HAES) in Aiyura, Eastern Highlands Province and at the Lowlands Agricultural Experiment Station (LAES) in Keravat, East New Britain.
- The International Potato Center collection (CIP) in Lima, Peru.
- The Vanuatu collection assembled in 2008 and conserved in the Agronomical and Technical Centre of Vanuatu in Santo (CTRAV), Vanuatu.
- The French Polynesia collection, collected in 2011 throughout the Marquesas, Society and Austral archipelagos and maintained in the research station of Papa’ara in Tahiti.
- Hawaiian landraces recovered from two botanical gardens, the University of Hawai’i Lyon Arboretum (HLA) (Oahu island) and the Waimea botanical garden (WBG) (Oahu island).

We also included some New Guinean landraces collected by Wulf Schiefenhövel (WS) in two Eipo villages of the West New Guinea highlands (Papua province of Indonesia). Healthy young leaves were collected from accessions maintained in greenhouses or fields and dried (except for Eipo material, which was conserved in salt) in an oven at 37 °C for two days or lyophilised, then conserved in silica gel.

Furthermore, we assembled a total of 57 herbarium specimens collected before 1978, of which 15 may be considered as modern (after the end of World War II) and 42 specimens from the 17th to the early 20<sup>th</sup> century, including historical specimens from the herbarium of Cook’s first voyage, conserved in the British Museum in London.

### *Genotyping*

All extraction and pre-PCR processing for herbarium specimens were conducted in a separate, isolated room, specifically designed for degraded DNA, with precautions taken to avoid and identify any potential contamination (38).

#### DNA extraction

For modern samples, DNA extraction was performed with the Qiagen 96 Plant kit for lyophilised tissues (Hilden, Germany). For herbarium specimens, extractions were performed from dry leaf tissues with the Qiagen DNeasy mini Plant Kit (250) (Hilden, Germany), with overnight lysis period. DNA was eluted in one time, in 100µl of TE. Specimens were extracted by sets of 12 samples, one set per day to prevent contaminations and two negative extraction controls were performed for each set of extractions.

#### PCR Amplification

Modern samples were genotyped at 11 nuclear microsatellite loci (J263, J315E, J522A, J116a, Ib297, J206A, J1809E, IbR16, IbC5, J544b, IbS11) and 4 chloroplast (cp) microsatellites described in Roullier et al. (2011). Two universal cp microsatellites (ccmp2 and ntcp28) used in Roullier et al. (2011) were replaced by new pairs of primers (ibcp4 F: ATCCTGGACGTGAAGAATAA, R: GATGGCTGAGTGGACTAAAG and ibcp31 F: AAACGGATTCTCCAATGTA, R: ACCTCACCGTTTCAGAAAGTA) specifically designed from the *Ipomoea purpurea* chloroplast genome, and giving more efficient amplification results. All loci were amplified independently using Multiplex PCR Taq (Qiagen) in a final volume of 10 µL, using 30ng of DNA per reaction. The following programme was conducted using a PTC-100 Thermocycler (MJ Research, Waltham, MA, USA): 15 min at 95°C , 35 cycles of 30 s at 94°C, 1 min 30 s at 57°C , 1 min at 72° C and finally 30 min at 72°C.

For herbarium specimens, the same chloroplast and nuclear loci were amplified independently, using Multiplex PCR Taq (Qiagen) in a final volume of 10 $\mu$ l, with 2  $\mu$ l of 1:10 DNA of unknown concentration, adding 0.1 $\mu$ l of 100X BSA. PCR program conditions were almost identical, but we performed in total 50 cycles. Because reproducibility problems are expected for the amplification of degraded DNA, we replicated each amplification. For haploid chloroplast loci, we amplified each individual two times independently. For hexaploid nuclear loci, we replicated each amplification between 2 and 5 times (depending on the priority given to the sample).

All amplification products were analyzed with an ABI 3130 XL 16-capillary sequencer (ABI Prism, Applied Biosystems, Foster City, CA, USA) and allele scoring was eye-checked by two investigators using Genemapper (Applied Biosystems, Foster City, CA, USA).

#### Herbarium specimens allele scoring

For chloroplast scoring, we conserved reproducible alleles (i.e., those present in both replicates) and retained only accessions with complete data. For nuclear loci, we used different scorings: the most stringent in which we kept alleles which were present at least three times over all replicates (maximum five); The second in which we retained alleles which were present at least two times. We included in the genetic analysis only accessions for which we obtained reliable information for at least five of the 11 loci. The number of missing data (missing loci) is given in Annexe 1. All herbarium specimens did not allow the same quality and reliability of scoring for nuclear genotyping.

#### *Assessing the genetic relationships between Oceanian and tropical American genepools with chloroplast and nuclear data*

We constructed the median joining networks of haplotypes from tropical America and Oceania using the software Network 4.5.6.1 (39) (<http://www.fluxusengineering.com.sharenet.htm>), and compared the proportions of haplotypes from each cp lineage in the two areas.

We then resorted to two kinds of methods to assign Oceania varieties to their neotropical source genepools on the basis of nuclear data. First, we used a non model-based method, the Discriminant Analysis of Principal Components (DAPC), a multivariate analysis recently developed and implemented in the Adegenet R package (40). We ran DAPC on the tropical American dataset, using Oceanian landraces as supplementary individuals, retaining five PCA components (29.3% of the total variance) for prior data transformation, for  $K = 2$  (prior grouping obtained following the  $K$ -means clustering algorithm implemented in the Adegenet package) (see Text S1 for details). We compared DAPC assignment with that obtained by a Bayesian model-based method implemented in the software Structure 2.3.3, recently adapted to handle polyploid data (41, 42). By pre-specifying source genepools in tropical America (as defined by the  $K$ -means clustering for  $K = 2$ ), the algorithm estimates ancestry for additional individuals (from Oceania), updating allele frequencies using only those from tropical America. We ran the admixture model with  $K = 2$  (same grouping as for DAPC), correlated allele frequencies, 50 000 burn-in iterations and 150 000 Markov chain-Monte Carlo steps and data coding for handling genotype ambiguity for codominant markers in polyploids (42).

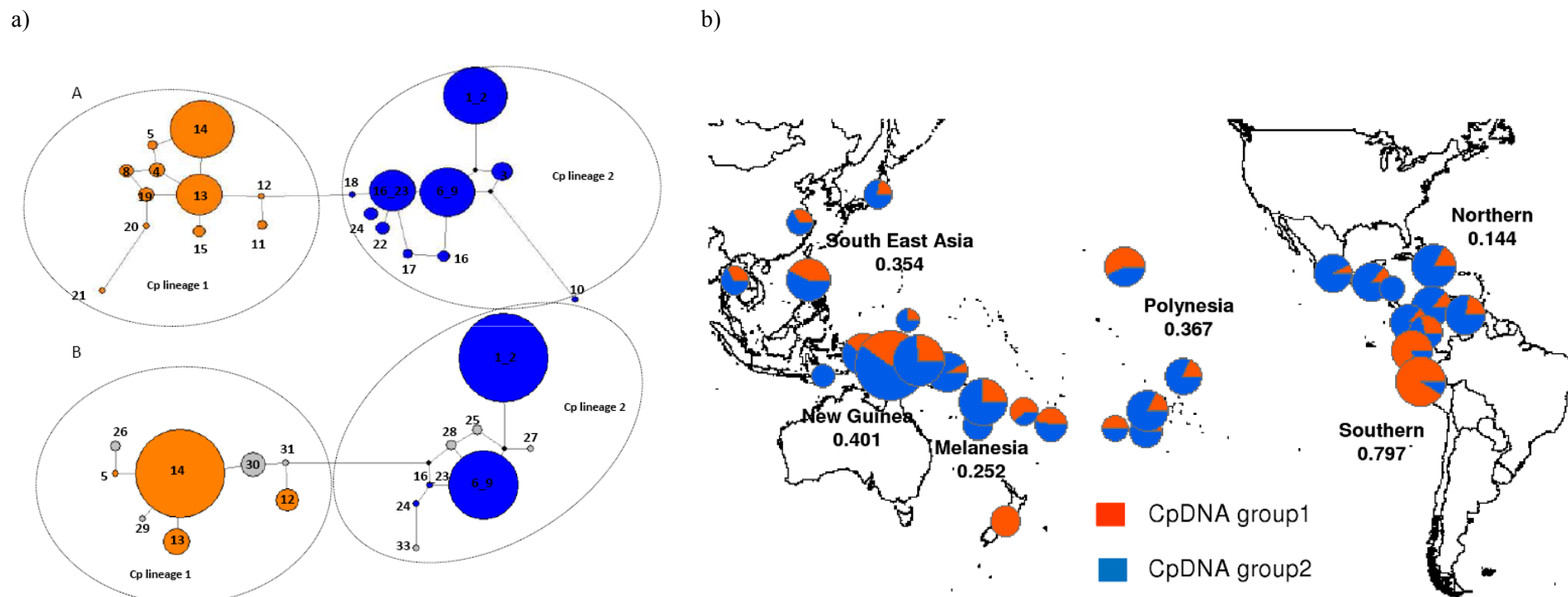
An individual is considered to be well assigned to its source genepool ( $K1$  or  $K2$ ) if the associated membership probability (DAPC) or ancestry value (Bayesian clustering) is greater than 0.8; otherwise the individual is considered “admixed” (AD).

Global genetic structure was also characterized by performing DAPC on the overall dataset, retaining 20 PCA components (54.5% of the total variance) for  $K = 12$  and constructing a Lynch distance-based Neighbour-joining tree (with the ape R package (43), then editing it with the Software Darwin 5.0.158 (<http://darwin.cirad.fr/darwin>). Classical diversity indices were computed using custom R scripts.

**Figure 6-2: Geographical distribution of nuclear and chloroplast genetic variation in sweet potato modern sampling.**

a) Comparison between median-joining networks obtained for tropical American (top) and Oceanian accessions (down). Circle size is proportional to the number of accessions that carry a given haplotype. Haplotypes from cp lineage 1 and 2 are colored in orange and blue, respectively. Private Oceanian haplotypes are indicated in gray. The length of the line between two distinct haplotypes is proportional to the number of mutations separating them. Sweet potato varieties in Oceania represent a subset of neotropical chloroplast diversity. They include representatives of both cp lineages 1 and 2 (in total seven different haplotypes, including those most common in tropical America) and also some private haplotypes (eight), all rare and derived from the most common ones by only one or two mutation steps.

b) Proportion of individuals belonging to each cp lineage (lineage 1 in orange, lineage 2 in blue) for each sampling site (country or archipelago). Only sites with four or more accessions were represented. Area of the circle is proportional to the square root of the sample size for the site. Values by region correspond to the frequency of cpDNA 1.



## RESULTS AND DISCUSSION

### *A summary of the situation in tropical America*

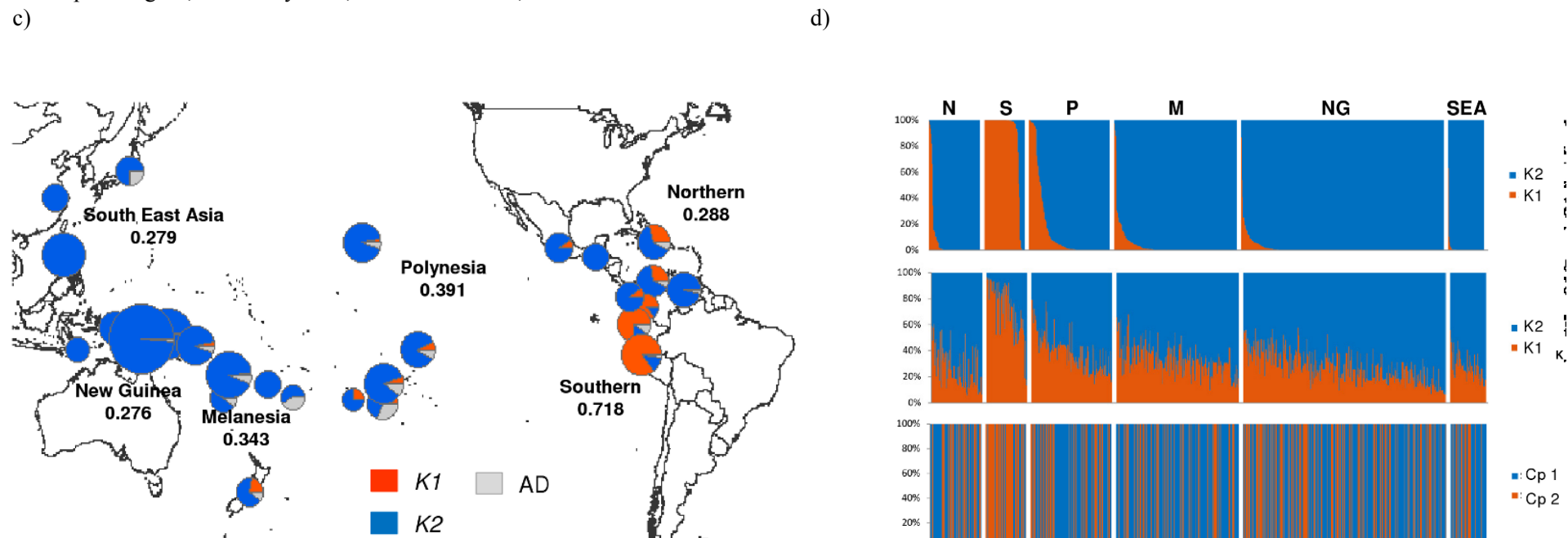
As discussed in previous studies (16, 17) and re-summarized here in Figures 6-2A, 6-2B, 6-2C and 6-2D, combining the use of chloroplast and nuclear microsatellite markers, we established the existence of distinct genepools in the Northern and Southern regions of the neotropics, certainly related to independent domestications. Neotropical sweet potatoes are characterized by two distinct geographically restricted chloroplast lineages (in total 21 haplotypes), which correspond quite well to two genetic clusters, *K1* and *K2*, identified by *K*-means clustering grouping: most accessions from the Southern region (hereafter the Southern genepool) exhibit chloroplast DNA (cpDNA) haplotypes of group 1 (79.7%) and belong to the nuclear cluster *K1* (83%). Sweet potatoes from the Northern region (hereafter the Northern genepool) carry mostly cpDNA haplotypes of group 2 (85.6%) and belong to the nuclear cluster *K2* (91.46%). Then, these genepools markers provide suitable tools to differentiate the Oceanian landraces originated from the Kumara line (derivating from the Southern genepool) from those originated from the Batata and Camote line (derivating from the northern genepool).

However, despite this clear-cut phylogeographic pattern, neotropical gene pools have secondarily come into contact, as shown by the admixture revealed with both chloroplast and nuclear markers: in the Southern region, we detected some accessions with cpDNA group 2 haplotypes (16/65), some assigned to the nuclear cluster *K2* (8/65 with the DAPC method and 3/65 with the Bayesian clustering method implemented in Structure 2.3.3), and some that were admixed individuals (3/65 with DAPC, 34/65 with the Bayesian method). Also, in the Northern region we identified several accessions with cpDNA group 1 haplotypes (31/82), some attributed to cluster *K1* (5/82 with DAPC) and some that were admixed individuals (2/82 with DAPC, 49/82 with the Bayesian method). This situation suggests that clones were exchanged between both regions and recombined with local material. While the DAPC method gives access to a more clear-cut phylogeographical pattern, the Bayesian method better underlines the initial situation of admixture between the two neotropical genepools.

**Figure 6-2:**

c) Proportion of individuals belonging to each nuclear cluster (cluster *K1* in orange, *K2* in blue) for each sampling site, as determined by the DAPC analysis. Only sites with four or more accessions were represented. Area of the circle is proportional to the square root of the sample size for the site. Values by region correspond to the mean *K1* ancestry as determined by the Bayesian clustering.

d) Three bar plots showing for each individual i) (top) the probabilities of membership in nuclear clusters *K1* and *K2* as determined by DAPC, ii) (middle), the probabilities of membership in nuclear clusters *K1* and *K2* as determined by the Bayesian clustering method implemented in Structure, and iii) (bottom), the individual's cp lineage. Each individual is represented as a vertical bar, with colors corresponding to membership probabilities in clusters *K1* (orange), *K2* (blue), and cp lineages 1 (orange) and 2 (blue). Individuals are organized by geographical origin, S for the Southern neotropical region, N for the Northern neotropical region, P for Polynesia, M for Melanesia, NG for New Guinea and SEA for South-East Asia.



Thus, admixture leads to some inconsistency between nuclear and chloroplast data which may complicate the assignation of Oceanian landraces to their original neotropical genepools.

*Assessing the relative contribution of Camote, Batata and Kumara lines of sweet potato introduction into Oceania in the modern sample*

Sweet potato varieties in Oceania represent a subset of neotropical diversity for both kinds of markers (Table 6-1). They include representatives of both neotropical cpDNA lineages 1 and 2 (in total eight different haplotypes, including those most common in tropical America) and also some private haplotypes (eight), all rare and derived from the most common ones by only one or two mutation steps (Figure 6-2A). Nuclear diversity indices (except for observed heterozygosity,  $H_o$ ) were slightly lower than those calculated for either of the two neotropical genepools (Table 6-1).

The neighbour-joining tree (and the DAPC analysis) revealed a high degree of phylogeographic structure among Oceanian varieties. As expected under the tripartite hypothesis, western Pacific varieties (varieties from South-East Asia, Melanesia and New Guinea) are differentiated from those of Polynesia (Figures 6-3A and 6-3B). However, incongruencies between chloroplast and nuclear data prevent in some cases the definitive attribution of some Oceanian varieties to their original genepool in tropical America.

Using the DAPC non model-based assignment method, most western Pacific sweet potato accessions (604 in total) were assigned to the cluster *K2*. Only four accessions (three from Melanesia, one from New Guinea) were assigned to the cluster *K1*, and 14 accessions exhibited a mixed genetic constitution (probability of membership in a given cluster  $< 0.8$ ) (Figures 6-2C and 6-2D). With the Bayesian clustering method, 144 accessions were assigned to the cluster *K2* and all others had a mixed ancestry. The mean proportion of *K1* ancestry was similar ( $0.29 \pm 0.128$ ) to that calculated in the Northern genepool ( $0.288 \pm 0.145$ ) (Figure 6-2C, Table 6-1). Thus, our analysis confirms previous results and indicates that the Northern genepool may have been the principal contributor to the nuclear genetic diversity of extant western Pacific sweet potatoes, which thus appear to be derived principally from the Camote and Batata lines. Furthermore, chloroplast lineage 2 haploypes were mainly represented throughout the western Pacific, as was the Northern genepool. However, the frequency of cpDNA lineage 1 was generally higher (33%) than that observed in the Northern genepool (14.4%) (Figure 6-2B and Table 6-1). This suggests that South American clones may have also been directly introduced throughout the western Pacific. Alternatively, founder effects related to diffusion history and/or local selection may have increased the frequency of this haplo-group in the area of introduction, compared to that found in the already admixed original Northern genepool.

**Table 6-1: Comparison of nuclear and chloroplast diversity between tropical America and Oceania.**

*N* corresponds to the total number of accessions from which nuclear, clonal and chloroplast diversity indices were calculated. *Na* is the mean number of alleles, *Ar* the allelic richness computed following a rarefaction procedure based on 1000 re-samplings of 50 accessions for each region, *He* the expected heterozygosity, *Ho* the observed heterozygosity, *R* the rarity index, and *KI* the mean value of *KI* ancestry provided by Bayesian clustering analysis. *MLG\_1* corresponds to the total number of distinct multilocus genotypes, and *MLG\_6* to the total number of multilocus genotypes with six or fewer pairwise differences. *H* is the total number of distinct haplotypes, and *cpDNA1* the proportion of cpDNA lineage 1 haplotypes.

Region	Chloroplast diversity			Nuclear diversity							Nuclear diversity		
	<i>N</i>	<i>H</i>	<i>cpDNA1</i>	<i>N</i>	<i>Na</i>	<i>Ar</i>	<i>He</i>	<i>Ho</i>	<i>R</i>	<i>KI</i>	<i>N</i>	<i>MLG_1</i>	<i>MLG_6</i>
Northern genepool	203	13	0.144	88	10	9.312	0.771	0.375	0.258	0.288	78	69	62
Southern genepool	152	15	0.797	69	9.182	8.746	0.729	0.335	0.217	0.718	55	54	49
Polynesia	148	6	0.367	131	8.636	7.33	0.76	0.384	0.132	0.391	96	83	54
Melanesia	214	7	0.252	209	8.636	7.204	0.76	0.388	0.116	0.343	159	136	113
New Guinea	364	10	0.401	329	9	7.702	0.74	0.383	0.096	0.276	267	229	192
South-East Asia	82	5	0.354	71	7.364	6.916	0.741	0.371	0.117	0.279	59	49	37
Oceania (total)	808	15	0.344	740	8.409	7.288	0.75	0.381	0.11	0.313	581	497	396



In Polynesia, the situation is more contrasted. In New Zealand, only cpDNA lineage 1 was found and in Hawaii this lineage dominated, a pattern expected for varieties that originated from the Southern genepool (Figure 6-2B). In contrast, the frequency of cpDNA lineage 1 was only 20% in eastern Polynesia (Figure 6-2B). Considering nuclear data, only 14 varieties (nine from eastern Polynesia, one from Hawaii and two from New Zealand) were assigned to cluster K1, 103 to cluster K2 and 14 had a mixed ancestry (DAPC non model-based assignment method) (Figure 6-2C). With the Bayesian clustering method, most of the accessions (127/131) had a mixed ancestry (Figures 6-2C and 6-2D). The mean proportion of K1 ancestry in Polynesia was 0.391 ( $\pm$  0.128) (Table 6-1) (ranging from 0.332 ( $\pm$  0.095) in Hawaii and 0.388 ( $\pm$  0.136) in eastern Polynesia to 0.585 ( $\pm$  0.186) in New Zealand), a value significantly higher than that calculated in the Northern genepool and the western Pacific ( $P < 0.01$  for both) attesting to the greater contribution of the Southern genepool in this area. However, only seven individuals exhibited a genetic constitution clearly inherited from the Southern genepool (cpDNA 1 lineage and assignment to cluster K1). Most modern accessions seem to have a mixed genetic constitution, “intermediate” between both genepools.

#### *Herbarium specimens and the Kumara line*

We recovered material from 42 herbarium accessions of sweet potato collected worldwide from the 17<sup>th</sup> century to the early 20<sup>th</sup> century, obtained data on chloroplast microsatellites for all of them and reliable data on nuclear microsatellites for 30 accessions. Almost all accessions from Polynesia exhibited a cpDNA lineage 1 haplotype (hap 14) and were assigned to nuclear cluster K1, a genetic background characteristic of the Southern genepool (Figure 6-4A). In contrast, herbarium specimens from the Caribbean, South-East Asia, and Madagascar carried a cpDNA lineage 2 haplotype (hap1\_2) and were assigned to nuclear cluster K2, a genetic constitution likely inherited from the Northern genepool. DAPC and Bayesian methods gave congruent results (Figure 6-4B). This clear geographical pattern provides strong support for the hypothesis that the Kumara line represents a pre-Columbian diffusion of sweet potatoes from South America (Peru-Ecuador area) into Polynesia.

Of salient interest are the specimens collected by J. Banks and D. Solander during Captain James Cook’s first voyage in 1769 (two plants from the Society Islands and one from New Zealand), because they may represent truly pre-historical introductions. All three derived from the Southern genepool (Figures 6-4A, 6-3A, 6-3B and 6-5). The two plants from the Society Islands are a single clone (multilocus genotypes almost identical). The New Zealand specimen, however, appears to be a distinct variety, as attested by morphological features and confirmed by our genetic results (incompletely genotyped, but still with distinct alleles). Other early 20<sup>th</sup> century specimens from eastern Polynesia also grouped with the Southern genepool (Figures 6-4A, 6-3A, 6-3B and 6-5). Morphologically distinguishable (annexe), they are genetically quite closely related (Figure 6-5) and may be true representatives of the Kumara line. This result confirms the conclusion of Green (20) that distinct and identifiable lineages were already present in Polynesia at the time of early European contacts, likely having however a narrow genetic base. Whether these lineages result from a single introduction from South America, or multiple independent introductions (possibly including voyages originating from northern Colombia and Central America) (27, 29), is still unknown and the present sample of herbarium accessions does not allow us to rule out multiple prehistoric introductions.

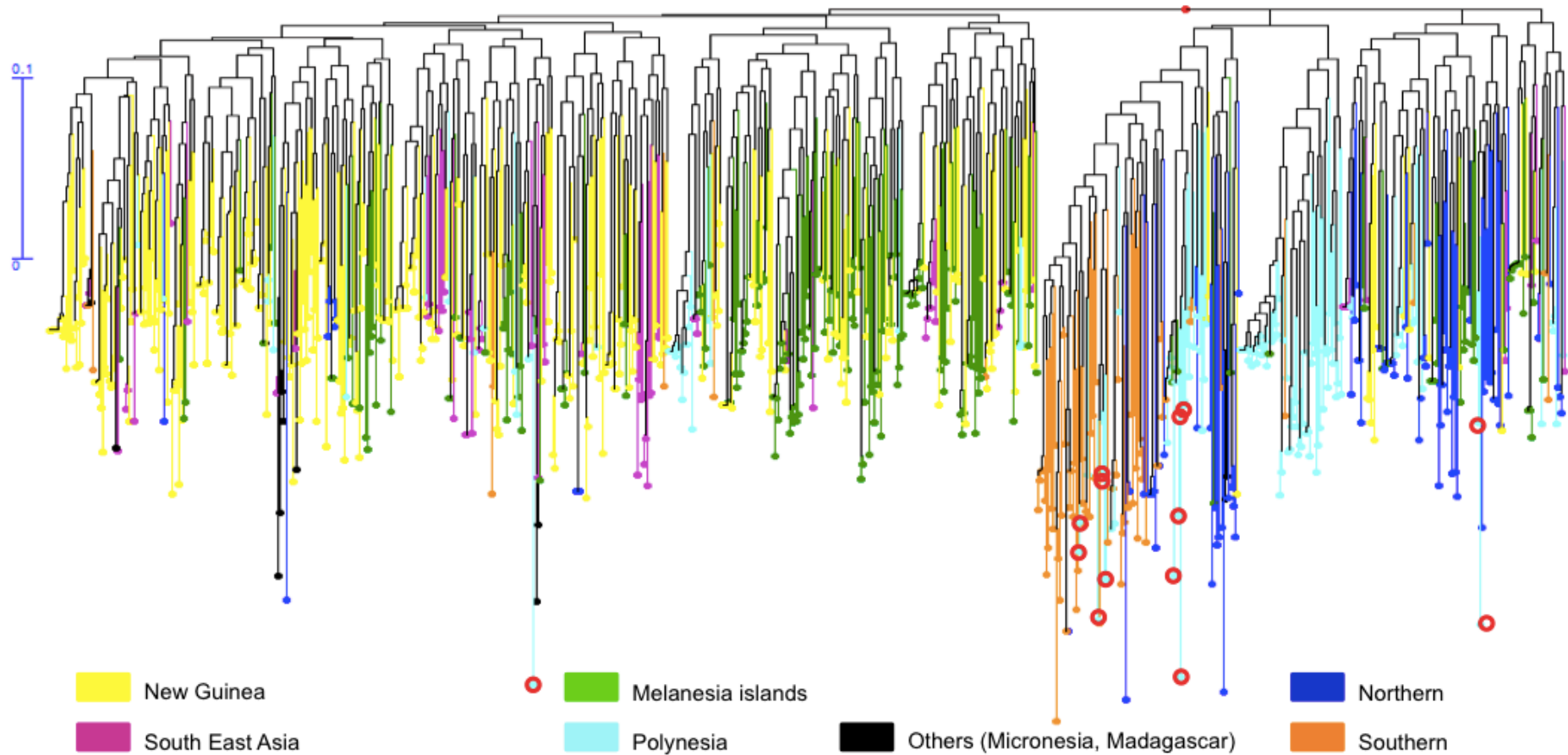
#### *From ancient to modern patterns: the reshuffling of the initial genetic background of sweet potato in Oceania*

##### **In Polynesia**

In contrast to all herbarium specimens up to the early 20<sup>th</sup> century, most of the contemporary varieties from eastern Polynesia (80.5%) carry a cpDNA lineage 2 haplotype, supporting the long-standing contention of ethnobotanists that later introductions have replaced initial sweet potato varieties in this area (19, 35, 44). The relative stability of genetic constitution reflected by herbarium accessions from the end of the 18<sup>th</sup> century to the beginning of the 20<sup>th</sup> century from this area may indicate that early European re-introductions (e.g., by traders, whalers and travelers (16)) did not immediately erase the initial genetic heritage. At first, these movements may have redistributed clones already present in the region rather than introducing new genotypes, extending the dispersal of varieties (which the Polynesians themselves had originated) throughout Polynesia and probably even further westward (18,

**Figure 6-3: Global patterns of genetic differentiation.**

a) Neighbour-joining tree based on the Lynch distance for the global dataset. Herbarium specimens from Polynesia are surrounded in red. Individuals are labelled according to their geographical origin, in orange, blue, magenta, green, yellow, violet and black for accessions from the Southern region, the Northern region, Polynesia, Melanesia, New Guinea, South-East Asia and “other regions” (Micronesia, Madagascar, Madeira), respectively.



19). The shift probably took place during the 20<sup>th</sup> century, surely before the 1960's, when Yen began his collections (19).

Differentiation patterns observed nowadays (Figure 6-2A, 6-2B and 6-5) reveal that varieties from eastern Polynesia form a heterogeneous group, with few accessions closely related to the Southern genepool, some to the Northern genepool and some to varieties predominant in the western Pacific, while some others are intermediate and even form a well-differentiated cluster (Figure 6-5). This group includes some clonal lineages, i.e. varieties with the same haplotype and only very few differences between genotypes, and also cultivars, although closely genetically related, which may be derived from independent sexual recombination events (Figures 6-5 and 6-6). This pattern of genetic differentiation suggests that later re-introductions did not simply replace initial ones but rather reshuffled the initial genetic background. Although mainly clonally propagated, sweet potato reproduces sexually, and its volunteer seedlings are sometimes incorporated by farmers and multiplied as new clones (19, 45). In the Marquesas and Tahiti, sweet potato long remained of only secondary importance. In the early 20th century, only a few native varieties (about five), poorly adapted to the humid tropical environments of eastern Polynesia, were recorded (46). Strong genetic bottlenecks, such as those that likely accompanied prehistoric diffusion of the crop by Polynesians who introduced a limited number of auto-incompatibility groups, may have greatly limited sexual reproduction and subsequent selection of seedlings by farmers as new clones. Later European re-introductions likely have allowed the initiation of local diversification. Hawaiian varieties exhibit a quite similar pattern to that shown by eastern Polynesian varieties and most of them are part of the eastern Polynesian cluster (Figure 6-5). Furthermore, some clones are shared between both regions (eastern Polynesia and Hawaii), attesting to the circulation of vegetative propagules between the two areas (Figure 6-5 and Table 6-2). Nevertheless, in Hawaii, early references by European explorers and traders described the sweet potato as a plentiful food (20). In the early 20th century, Handy (47) characterized a great number of local varieties with a wide diversity of uses. Archeological data revealed that sweet potato was probably the dominant crop as early as 1400 A.D. in the drier areas on the leeward side of the archipelago (48). It is possible that independent pre-historical introductions or early historical introductions by Spanish explorers (Figure 6-1) broadened the initial genetic base, thus providing quite rapidly opportunities for widespread recombination and local selection of variants in this archipelago. Interestingly, early 20<sup>th</sup> Century herbarium specimens already exhibit a "mixed" constitution (Figures 6-3A and 6-3B).

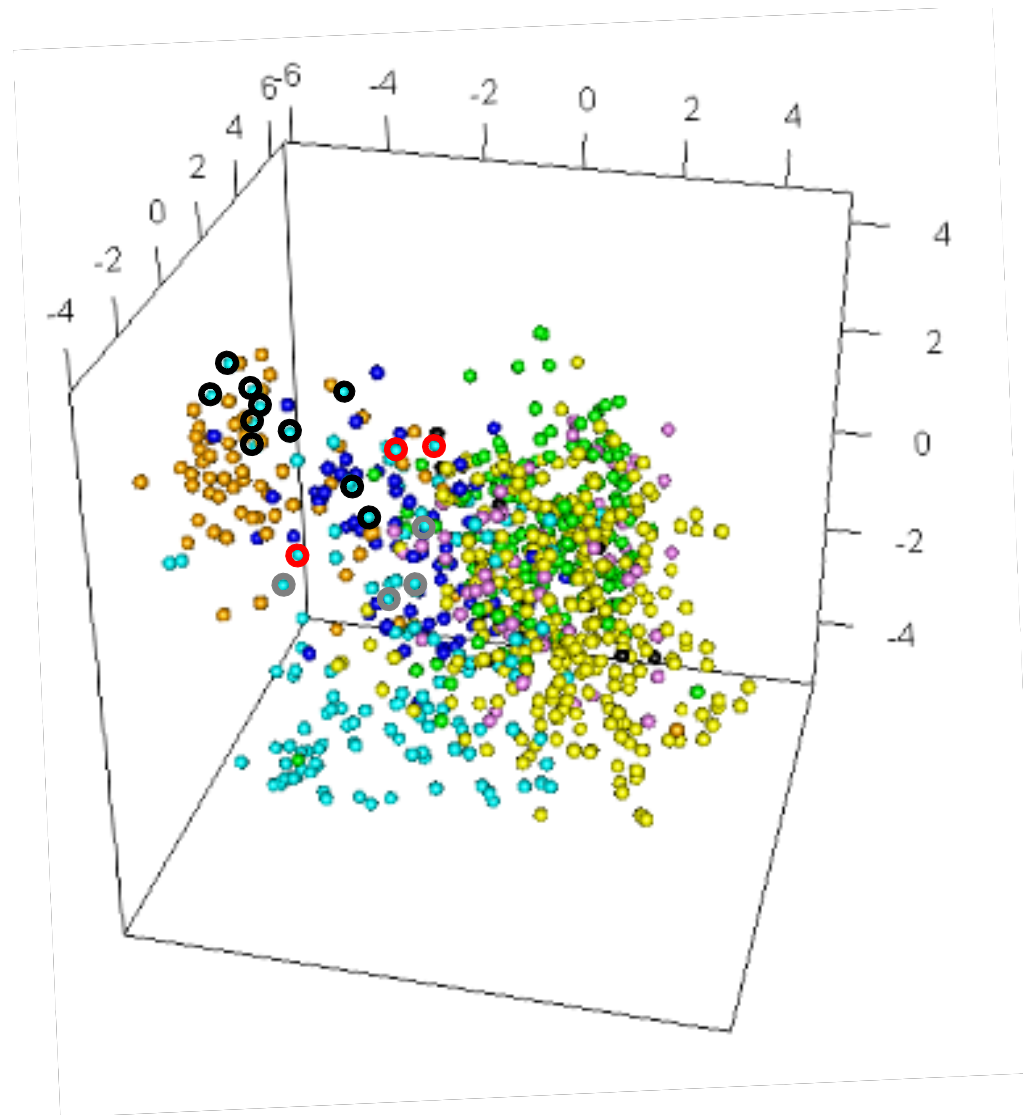
None of the New Zealand varieties group with the eastern Polynesia/Hawaii cluster (Figure 6-5). Two cultivars were clearly associated with the Southern genepool on the basis of both kinds of markers and may represent true original Kumaras. Waina varieties, the main commercial sweet potato cultivars in New Zealand, cluster together, even forming a clonal lineage. This is consistent with historical data which claimed that Waina are derived from a single variety introduced by whalers who visited New Zealand aboard the *Rainbow* in 1860 (49). Two other accessions, recognized by Maori informants to be native Kumaras, i.e., pre-European introductions, are associated with the Northern genepool and may instead represent later re-introductions.

The modern genetic diversity in Polynesia thus results from the local reshuffling of initial genetic background with later re-introductions. The observed pattern also reflects the likelihood that Polynesia remained relatively isolated from the rest of the Pacific. Movements of sweet potato germplasm between the eastern and western Pacific were quite restricted and did not homogenize sweet potato genetic diversity in Oceania. Indeed, very few clones are shared between Polynesia and other parts of Oceania (Table 6-2).

### **In the western Pacific**

In contrast, the lack of genetic differentiation among western Pacific varieties (Figures 6-3A and 6-3B) suggests a common genetic background consistent with the tripartite hypothesis. Moreover, this pattern suggests a wide circulation of landraces across this region, also attested to by the presence of several shared clones (Table 6-2). Still, the mean value of *K1* ancestry found in Melanesia was slightly higher ( $0.343 \pm 0.113$ ) than that observed in New Guinea ( $0.276 \pm 0.116$ ) and in South-East Asia ( $0.279 \pm 0.097$ ), which both showed values very similar to that observed in the Northern genepool. The greater contribution of the Southern genepool in Melanesia likely reflects early introductions of true Kumara clones, as also suggested by the common use of cognates of this term ("kumara/kumala")

b) Scatter plot of DAPC (first three discriminant axes) performed on the global dataset for  $K = 12$ . Individuals are labelled with the same colors than for the tree. Herbarium specimens from eastern Polynesia, Hawaii and New Zealand are represented by symbols surrounded in black, gray and red, respectively



to designate sweet potato in eastern Melanesia (19, 35). Also, New Guinea varieties appear to be slightly differentiated from those of other areas in the western Pacific (Figures 6-2A and 6-2B). These genetic data, combined with the presence of many seemingly unrelated names, likely reflect multiple processes of local diversification in this island. Also, the frequency distribution of pairwise Manhattan distances for western Pacific genotypes (see also chapter 5) suggest that the high extent of diversity in the western Pacific appears to have resulted from intensive recombination among introductions and from the use by farmers of plants issued from true seed (Figure 6-6) rather than from clonal evolution and selection by farmers of somatic mutants, as suspected by Yen (1974).

The history of the sweet potato in Oceania thus demonstrates how plant movements and genetic recombination between initially differentiated genepools have continuously reshuffled the crop's genetic background, providing farmers the raw material to generate quite rapidly an impressive number of variants adapted to a wide range of environments.

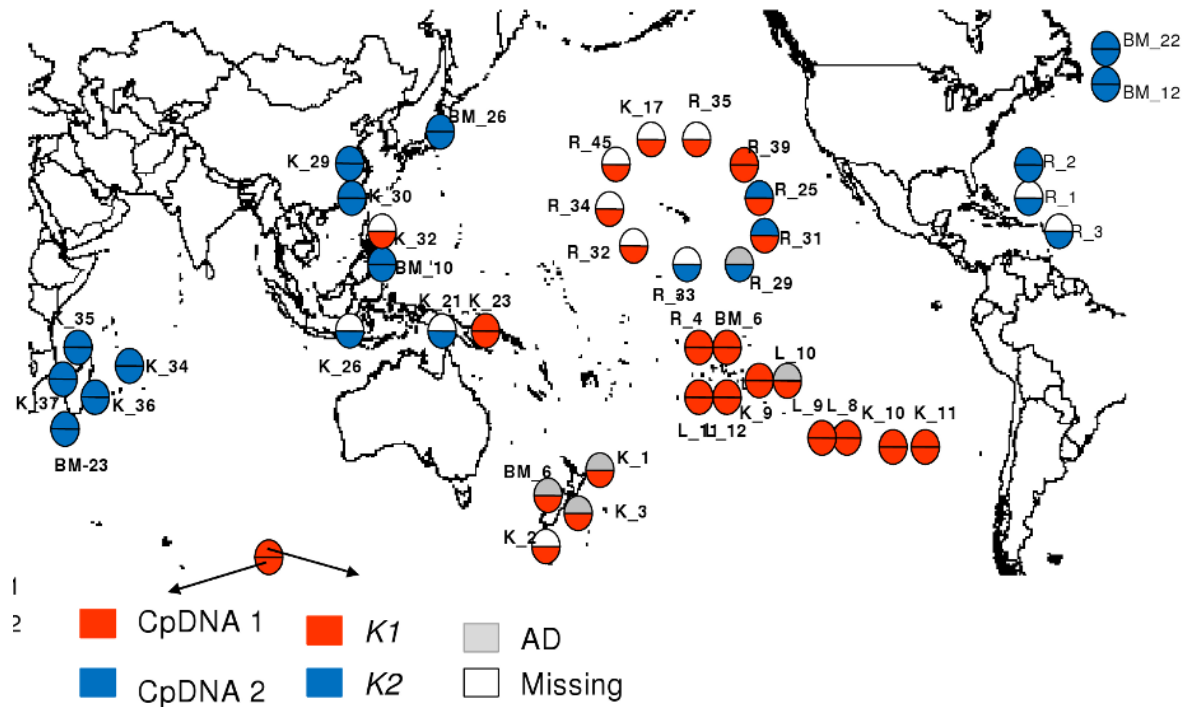
#### *Did genes and names disperse together?*

A recent multidisciplinary study on banana proposed a complete reconstruction of scenarios for the origin and dispersal of banana cultivars supported by both linguistic and genetic data (8). In this particular case, where clonal propagation has long been almost exclusive (seedless fruits preclude future genetic recombinations), names and genotypes share similar geographical and historical associations. Over successive dispersal events, names accumulate traceable terminological changes, straightforwardly associated with the genotype, itself relatively fixed by vegetative propagation.

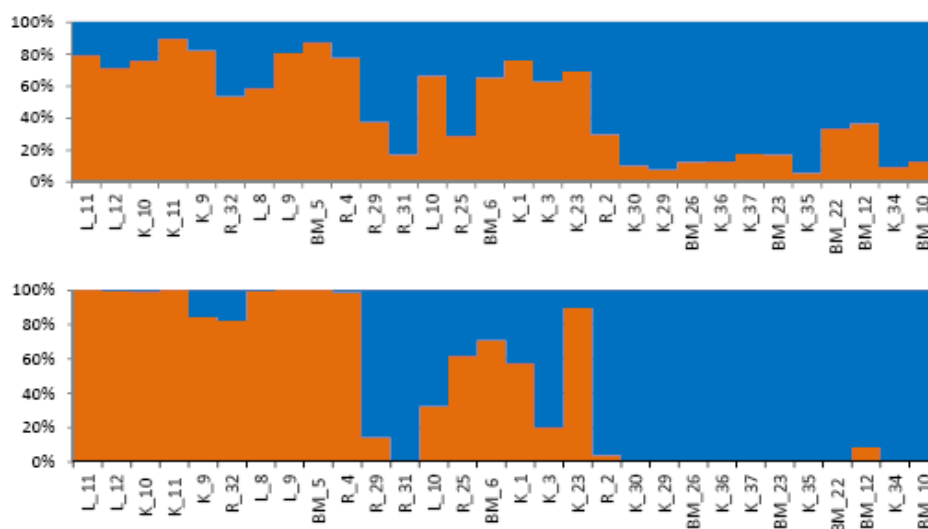
In contrast, sweet potato exhibits a mixed clonal/sexual reproductive system. As in banana, names are likely modified when the plant moves. Furthermore, in sweet potato newly introduced genotypes should often recombine with local material. In these recombined genotypes, the name likely follows the phenotype. Distinctive phenotypes may reflect alleles at loci that are under strong selection, and may show no correlation with the neutral genetic background, widely reshuffled. Thus, old names may continue to be applied to landraces that have been strongly affected by modern plant movement and local recombination. This is what likely happened to most of the “native” Polynesian varieties originating from South America, which were progressively admixed with material introduced much later. An illustration of this evolutionary scenario is offered by the group of ancient Polynesian sweet potatoes categorized as *Convolvulus chrysorrhizus dulcis* by D. Solander and subsequently by J. Forster, because of the bright yellow color of their internal flesh. In 1960, Yen identified two very similar groups of yellow varieties (yellow internal flesh and fusiform root) with a pan-Polynesian distribution (the “re’amoā” and “hererai” groups) and considered by local people to be native Kumaras (19). While the “native” yellow variety described by Banks and Solander carries a cpDNA lineage 1 haplotype, all the contemporary yellow varieties bear a cpDNA lineage 2 haplotype, and are very closely related, probably forming a clonal lineage or a group of genetically closely related varieties, confirming the suspicions of Yen (19). The neutral genetic basis has shifted, while the initial phenotype, and probably the initial names, may have been well conserved.

**Figure 6-4: Genetic constitution of herbarium specimens collected from the 18<sup>th</sup> century to the early 20<sup>th</sup>**

a) Top and bottom halves of each circle represent the cp lineage and nuclear cluster (as determined by DAPC), respectively.



b) Bar plots showing a) (top) the probabilities of membership in nuclear clusters K1 (orange) and K2 (blue) as determined by DAPC method, b) (bottom), the ancestry value in nuclear clusters K1 (orange) and K2 (blue) as determined by the Bayesian method, for old herbarium specimens (from the 17<sup>th</sup> to early 20<sup>th</sup> century).

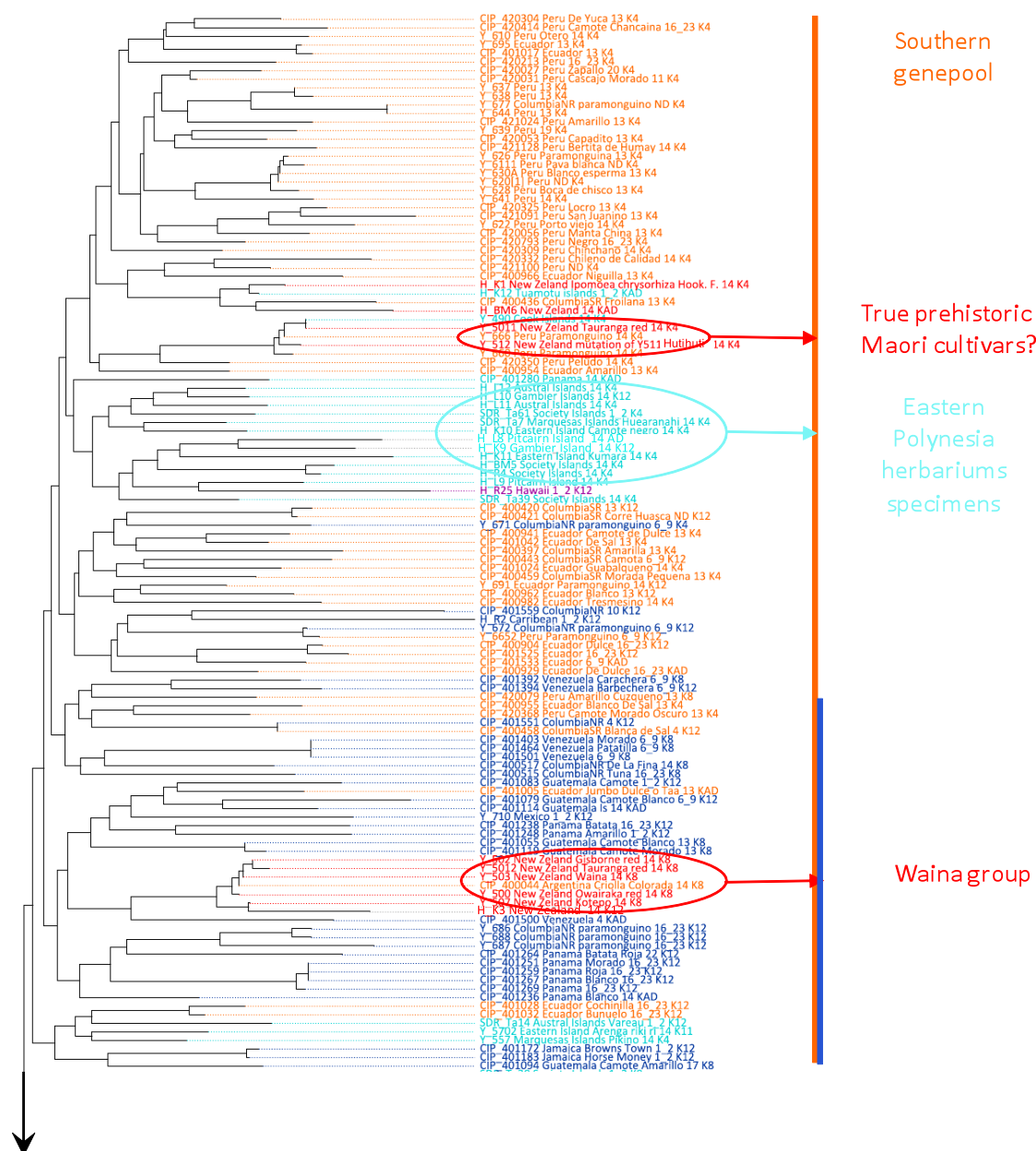


The same scenario likely occurred for Maori cultivars. Three cultivars are recognized by Maori informants to be native Kumaras, i.e., pre-European introductions (19, 20, 49). They are “Rekamoroa”, “Taputini” and “Hutihuti” cultivars. A previous genetic study using RAPD markers found “Rekamoroa” and “Hutihuti” to be very closely related and “Taputini” quite distinct (50), while Yen suggested on morphological grounds that “Hutihuti” and “Taputini” were closely related (19). A recent genetic study based on AFLP showed that none of these cultivars group with other eastern-Polynesian varieties (35). The author of this last study concluded that neither of these cultivars may be derived from true pre-historic introductions but instead from European introductions, like the majority of New Zealand commercial cultivars (the Waina group). We found similar results: the Waina group and the “Rekamoroa” and “Taputini” cultivars grouped with accessions from the Northern genepool, and not with the group of accessions from eastern Polynesia, suggesting that they derived from latter European introductions. However, accessions of the cultivar “Hutihuti” grouped with accessions from the Southern genepool, as did the historical herbarium specimens from eastern Polynesia. Only this last cultivar may represent a true native Kumara. However, Clarke (35) revealed some problems of misidentification between independent collections (same cultivar with different names in the two collections) in Japan and New Zealand, which complicate the interpretation of genetic data. Thus, before conclusions can be reached about this debated topic, investigations must first verify the true identities (names) of the plants concerned, crossing information in Yen (19) and data on agromorphological traits of living plants.

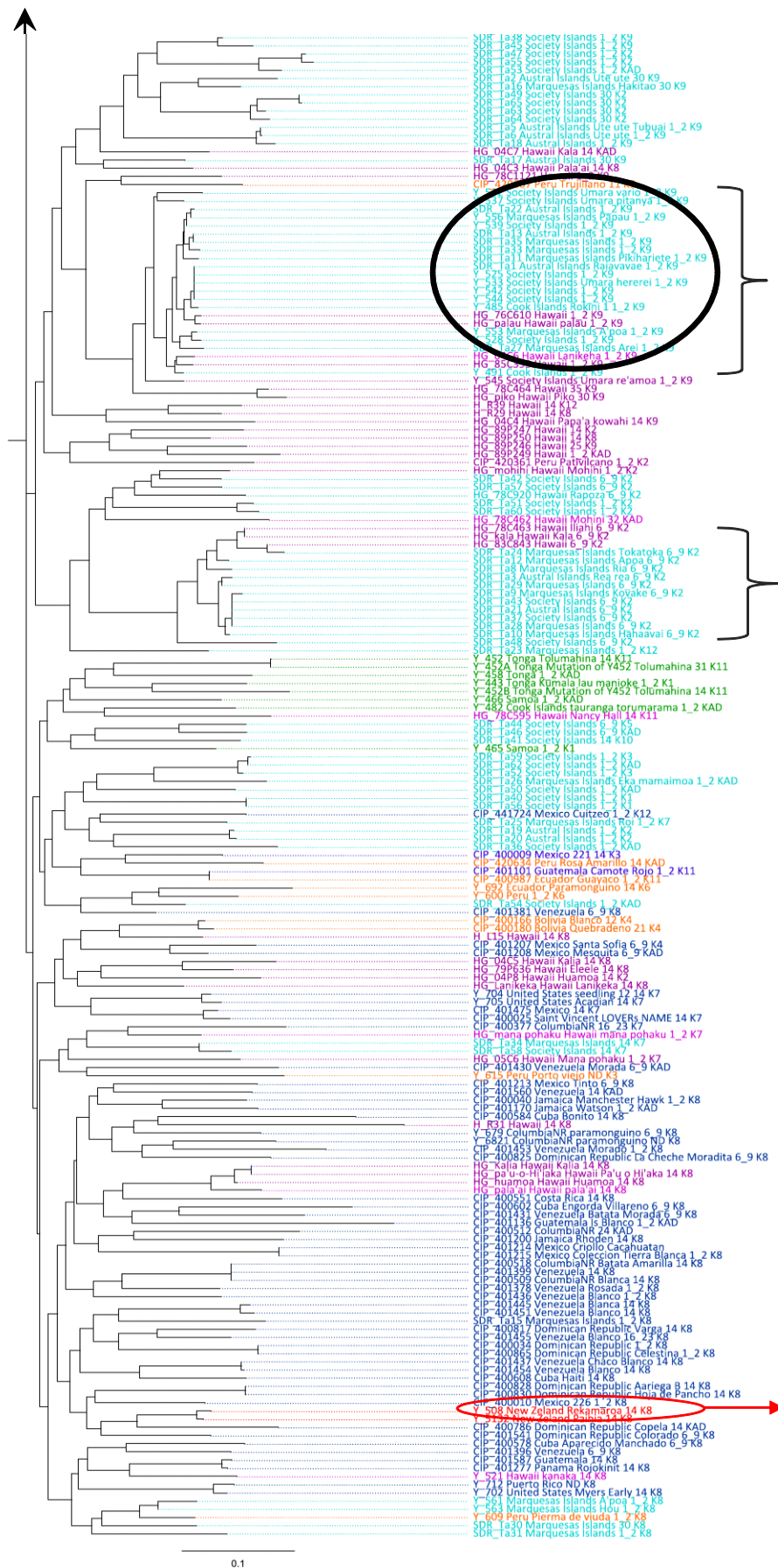
Patterns in eastern Melanesia may have a similar explanation. The common name for sweet potato in this region is derived from the Polynesian term “kumara” or “kumala”, suggesting early sweet potato transfers from Polynesia (by Polynesians themselves, by early Europeans, or both) (19, 30). However, the pattern of variation for neutral nuclear markers is similar to that found in western Melanesia, South-East Asia and New Guinea, indicating that if true Kumaras were initially introduced, they have been widely replaced by, or reshuffled with, Camote/Batata varieties.

Thus genotypes and names do not always migrate together, and information on each gives access to complementary parts of the plant’s dispersal history. Whereas old names and old phenotypes may endure, recombination may widely reshuffle their neutral genetic backgrounds and obscure the pathways of initial diffusion.

**Figure 6-5: Hierarchical representation of the radial tree** showing only accessions from the Southern region (in orange), Northern region (in blue), and Polynesia (Eastern Polynesia in magenta, Hawai'i in violet, New Zealand in red and Western Polynesia (Tonga) in green). Accession labels comprise the collection name (Y for Yen collection, CIP for the International Potato Centre collection, HG for Hawai'i botanical Garden accessions, H for herbarium specimens and SDR for the French Polynesian collection maintained in the rural development service), the sample number (referred to in Table S1), the geographical origin (country or archipelago), the local name if determined, the cp DNA haplotype and the associated nuclear cluster as determined by the global DAPC analysis (for  $K = 12$ ).







Eastern  
Polynesia/Hawaii  
cluster

Yen's "Re'amoia"  
and "Hererei"  
lineages

Northern  
genepool

New Zealand  
«Rekamaroa»  
and «Taputini»  
varieties

## Conclusion

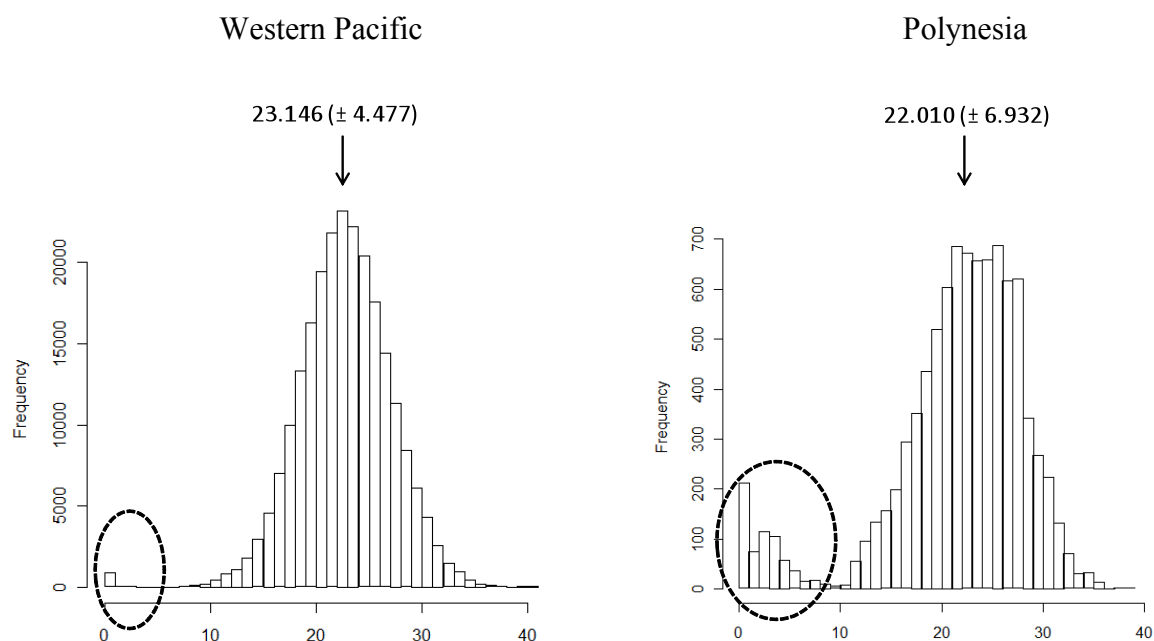
Our study has shown how historical collections reveal patterns of diffusion of sweet potato in Oceania that have been obscured by modern plant movements and local recombination. Our results provide strong genetic support, previously lacking, for the tripartite hypothesis, notably concerning the Kumara line, the pre-Columbian diffusion of sweet potatoes from South America into Polynesia. We also suggest why phenotypes, names and neutral genes do not necessarily share completely parallel evolutionary histories. Exploiting crops as a proxy of human movements in Oceania requires multidisciplinary approaches combining linguistics, morphological characterization of plants and phylogeography, in which the use of herbarium collections may play a critical role.

## Acknowledgments

This study was funded by grants from the “Pacific funds” distributed by the French Ministry of European and Foreign Affairs and received support from the SYNTHESYS Project (<http://www.synthesys.info/>), which is financed by European Community Research Infrastructure Action under the FP7 “Capacities” Program. CR was supported by a doctoral grant from the French Ministry of Higher Education and Research. The authors are very grateful to Marc Spencer, Jonathan Gregson and Jacek Wajer (Natural History Museum, London, England), David Goyder and Felix Forest (Royal Botanical Gardens, Kew, England), Gerard Thijssse (National Herbarium Nederland, Leiden University branch, Netherlands) and Napua Harbottle (Bishop Museum, Honolulu, Hawaii, USA) for providing herbarium specimens. We also thank Toru Kumagai (National Institute for Agronomical Sciences, Tsukuba, Japan), Maurice Wong (Rural Development Services, Papeete, Tahiti) and Wulf Schiefenhövel (Max-Planck-Institute for Ornithology) for their help in the sampling of “modern” sweet potato material. Data used in this work were produced through the technical facilities of the Mediterranean Center for Environment and Biodiversity.

**Figure 6-6: Frequency distribution of genetic dissimilarity based on the calculation of all pairwise Manhattan distances between genotypes within landraces of Western Pacific and Polynesia.**

The first peaks of the bimodal curve represent clones (pairwise dissimilarities equal to 0), possible variants by somatic mutations and/or scoring errors. The second part of the curve represent individuals which may derivate from independent sexual recombination events. The mean value of pairwise distance and its standard deviation is provided for each region.



## References

1. Thorsby E (2012) The Polynesian gene pool: an early contribution by Amerindians to Easter Island. *Proc R Soc Lond B-Biol Sci* 367:812-819.
2. Kayser M (2010) The human genetic history of Oceania: near and remote views of dispersal. *Curr Biol* 20:194-201.
3. Kirch PV (2000) *On the Road of the Winds: An Archaeological History of the Pacific Islands Before European Contact* (University of California Press, Berkeley)
4. Matisoo-Smith E et al. (1998) Patterns of prehistoric human mobility in Polynesia indicated by mtDNA from the Pacific rat. *Proc Natl Acad Sci USA* 95:15145-15150.
5. Matisoo-Smith E, Robins JH (2004) Origins and dispersals of Pacific peoples: evidence from mtDNA phylogenies of the Pacific rat. *Proc Natl Acad Sci USA* 101:9167-9172.
6. Erickson DL, Smith BD, Clarke AC, Sandweiss DH, Tuross N (2005) An Asian origin for a 10 000-year-old domesticated plant in the Americas. *Proc Natl Acad Sci USA* 102:18315-18320.
7. Clarke AC, Burtenshaw MK, McLenachan PA, Erickson DL, Penny D (2006) Reconstructing the origins and dispersal of the Polynesian bottle gourd (*Lagenaria siceraria*). *Mol Biol Evol* 23:893-900.
8. Perrier X et al. (2011) Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proc Natl Acad Sci USA* 108:11311-11318.
9. Storey A et al. (2007) Radiocarbon and DNA evidence for a pre-Columbian introduction of Polynesian chickens to Chile. *Proc Natl Acad Sci USA* 104:10335-10339.
10. Gongora J et al. (2008) Indo-European and Asian origins for Chilean and Pacific chickens revealed by mtDNA. *Proc Natl Acad Sci USA* 105:10308-10313.
11. Larson G et al. (2007) Phylogeny and ancient DNA of *Sus* provides insights into Neolithic expansion in Island Southeast Asia and Oceania. *Proc Natl Acad Sci USA* 104:4834-4839.
12. Moodley Y et al. (2009) The peopling of the Pacific from a bacterial perspective. *Science* 323:527-530.
13. Lawler A (2010) Beyond Kon-Tiki: did Polynesians sail to South America? *Science* 328:1344-1347.
14. Jones TL, Storey AA, Matisoo-Smith E, Aliaga Ramirez JM (2010) *Polynesians in America: Pre-Columbian contacts with the New World* (Alta Mira, Plymouth).
15. Austin DF (1988) in *Exploration, maintenance, utilization of sweet potato genetic resources* (Proceedings of the first Planning Conference, Lima), pp 27-59.
16. Roullier C, Rossel G, Tay D, McKey D, Lebot V (2011) Combining chloroplast and nuclear microsatellites to investigate origin and dispersal of New World sweet potato landraces. *Mol Ecol* 20:3963-3977.
17. Roullier C et al. Disentangling the origins of cultivated sweet potato. Submitted to PLoS ONE.
18. Barrau J (1957) L'énigme de la patate douce en Océanie. *Etudes d'Outre-Mer* 40:83-87.
19. Yen DE (1974) *The Sweet Potato in Oceania: An Essay in Ethnobotany* (Bishop Museum Press, Honolulu).
20. Green R. (2005) in *The Sweet Potato in Oceania: a Reappraisal*, eds Ballard C, Brown P, Bourke RM, Harwood T (Oceania Publications, Sydney), pp 43-62.
21. Hather JG, Kirch PV (1991) Prehistoric sweet potato (*Ipomoea batatas*) from Mangaia Island, central Polynesia. *Antiquity* 65:887-893.
22. Ladefoged TN, Graves MW, Coil J (2005) The introduction of sweet potato in Polynesia : early remains in Hawaii. *J Polyn Soc* 114:359-374.
23. Horrocks M, Rechtman RB (2009) Sweet potato (*Ipomoea batatas*) and banana (*Musa* sp.) microfossils in deposits from the Kona Field System, island of Hawaii. *J Archaeol Sci* 36:1115-1126.

**Table 6-2: Multilocus clonal lineages shared between regions for the modern sample** (multilocus genotypes which have six or fewer pairwise differences and may be clones or genetically closely related individuals).

Numbers refer to those identifying these lineages in annexe 1a. ND means not determined.

	Clonal lineage	1	2	3	4	5	14	18	25	32	34	35	37	40	49	50	57	66	67	70	72	78	93	96
Region	Haplotype	1_2	14	1_2	1_2	1_2/6_9	14	1_2	14	1_2	14/ND	1_2/ND	14/ND	6_9	6_9	14	1_2	1_2	14	1_2	6_9	6_9	13/ND	6_9
Northern region	Colombia																				1		1	
	Guatemala															1	1							
	Panama															1								
Southern region	Argentina		1																					
	Ecuador																1							
	Peru										1	1							2	1	1		1	
Eastern Polynesia	Austral Islands	1																				2		
	Cook Islands	2																	1					
	Marquesas Islands	5																		1		3		
	Society Islands	6																				2		
	Hawaii	3																				2		
	New Zealand		3																2					
Western Polynesia	Tonga																	1						
Micronesia	Caroline Islands											1												
Eastern Melanesia	Fiji	1														1		1						
	New Caledonia	1																						
	Vanuatu		1	2		1	1								2	1								1
Western Melanesia	Bismarck arch.				8	1	1		4				1	1										
	Solomon Islands							1	1				2	2										1
New Guinea	New Guinea			1	1			2		1	1													
South-East Asia	China											2												
	Japan											1			1									
	Timor								2															
	<b>Total</b>	<b>19</b>	<b>5</b>	<b>3</b>	<b>9</b>	<b>2</b>	<b>2</b>	<b>3</b>	<b>5</b>	<b>3</b>	<b>2</b>	<b>5</b>	<b>3</b>	<b>3</b>	<b>3</b>	<b>4</b>	<b>2</b>	<b>2</b>	<b>5</b>	<b>2</b>	<b>2</b>	<b>9</b>	<b>2</b>	<b>2</b>

24. Wallin P, Stevenson C, Ladefoged T (2005) in *The Sweet Potato in Oceania: a Reappraisal*, eds Ballard C, Brown P, Bourke RM, Harwood T (Oceania Publications, Sydney), pp 85-88.
25. Horrocks M, Shane PA, Barber IG, D'Costa DM, Nichol SL (2004) Microbotanical remains reveal Polynesian agriculture and mixed cropping in early New Zealand. *Rev Palaeobot Palynol* 131:147-157.
26. Bulmer R (1966) Birds as possible agents in the propagation of the sweet potato. *The Emu* 65:165-182.
27. Montenegro Á, Avis C, Weaver A (2008) Modeling the prehistoric arrival of the sweet potato in Polynesia. *J Archaeol Sci* 35:355-367.
28. Scaglione R (2005) in *The Sweet Potato in Oceania: a Reappraisal*, eds Ballard C, Paula B, Bourke RM, Harwood T (Oceania Publications, Sydney), pp 35-42.
29. Rensch KH (1991) in *Polynesian Plant Names: Linguistic Analysis and Ethnobotany, Expectations and Limitations*, eds Cox P, Banack S (Dioscorides Press, Portland), pp 97-111.
30. Allen MG (2005) in *The Sweet Potato in Oceania: a Reappraisal*, eds Ballard C, Brown P, Bourke RM, Harwood T (Oceania Publications, Sydney), pp 99-108.
31. Baert A MG (2001) in *Histoire de la découverte des régions australes: les Salomon, Marquises, Santa Cruz, Tuamotu, Cook du Nord et Vanuatu* (l'Harmattan, Paris).
32. Zhang D, Ghislain M, Huam Z, Golmirzaie A, Hijmans R (1998) RAPD variation in sweetpotato (*Ipomoea batatas* (L.) Lam.) cultivars from South America and Papua New Guinea. *Genet Resour Crop Evol* 45:271-277.
33. Zhang D, Rossel G, Kriegner A, Hijmans R (2004) AFLP assessment of diversity in sweetpotato from Latin America and the Pacific region: Its implications on the dispersal of the crop. *Genet Resour Crop Evol* 51:115-120.
34. Gichuki ST et al. (2003) Genetic diversity in sweetpotato [*Ipomoea batatas* (L.) Lam.] in relationship to geographic sources as assessed with RAPD markers. *Genet Resour Crop Evol* 50:429-437.
35. Clarke A (2009) *Origins and Dispersal of the Sweet Potato and Bottle Gourd in Oceania: Implications for Prehistoric Human Mobility*. PhD thesis (Massey University, Palmerston North, New Zealand).
36. Bourke (2005) in *The Sweet Potato in Oceania: a Reappraisal*, eds Ballard C, Paula B, Bourke RM, Harwood T (Oceania Publications, Sydney), pp 171-180.
37. Matisoo-Smith E (2002) Something old, something new: do genetic studies of contemporary populations reliably represent prehistoric populations of Pacific *Rattus exulans*? *Hum Biol* 74:489-496.
38. Knapp M, Clarke AC, Horsburgh KA, Matisoo-Smith E (2012) Setting the stage - building and working in an ancient DNA laboratory. *Ann Anat* 194:3-6.
39. Bandelt HJ, Forster P Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37-48.
40. Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11: 94.
41. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.
42. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164:1567-1587.
43. Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289-290.
44. Barrau J (1961) *Subsistence Agriculture in Polynesia and Micronesia* (Bernice P. Bishop Museum, Honolulu).
45. Schneider J (1995) in *Indigenous Knowledge in Conservation of Crop Genetic Resources*, ed Schneider J (Proceedings of an International Workshop, Bogor), pp 63-70.
46. Brown FBH *Flora of Southeastern Polynesia, Volume III* (Bernice P. Bishop Museum, Honolulu).

47. Handy ESC (1940) *The Hawaiian Planter. Volume I: His Plants, Methods and Areas of Cultivation.* (Bernice P. Bishop Museum, Honolulu).
48. Coil J, Kirch PV (2005) in *The Sweet Potato in Oceania: a Reappraisal*, eds Ballard C, Brown Paula, Bourke RM, Harwood T (Oceania Publications, Sydney), pp 71-84.
49. Stokes, J. (2007, February 1). Kumara claim becomes hot potato. *New Zealand Herald*. Retrieved June 21, 2010, from [http://www.nzherald.co.nz/jonstokes/news/article.cfm?a\\_id=96&objectid=10421870](http://www.nzherald.co.nz/jonstokes/news/article.cfm?a_id=96&objectid=10421870)
50. Harvey C, Gill G, Crossman C, Fraser, L (1997). Assessing relationships of kumara cultivars by RAPD analysis. *New Zealand Journal of Botany* 35: 479-485.





# *Chapitre 7*

## *Discussion générale*

# I. AUTO OU ALLOPOLYPLOIDIE ? DE LA DIFFICULTE DE DEFINIR LA NATURE DES POLYPLOÏDES

Les polyploïdes sont classiquement regroupés en deux classes, les autopolyploïdes et les allopolyploïdes (Ramsey et Schemske 1998, 2002). Deux types de définitions sont fréquemment avancés. Selon la définition taxonomique, les autopolyploïdes ont évolué à partir d'une seule espèce alors que les allopolyploïdes ont évolué à partir de plusieurs espèces. Selon la définition cytologique, les autopolyploïdes sont issus d'individus dont les parents forment des hybrides fertiles et sont caractérisés par la formation de multivalents à la méiose (Stebbins 1950, 1980 ; Levin 2002). De l'autre côté, les allopolyploïdes sont issus d'un croisement entre individus dont l'hybride a une fertilité réduite (Stebbins 1950, 1980; Levin 2002). Le dédoublement du génome est nécessaire pour restaurer la fertilité de l'hybride et l'allopolyploïde ainsi formé sera caractérisé par la formation de bivalents à la méiose, où chaque chromosome s'apparie avec le chromosome provenant du même individu parental. Selon cette définition, les allopolyploïdes devraient préserver des traces de leurs deux parents puisque les chromosomes provenant de chacun ne ségrègent pas et sont transmis de génération en génération. Les autopolyploïdes forment eux des multivalents où tous les chromosomes peuvent ségréger à la méiose. La trace de l'origine des parents peut donc être perdue rapidement.

Ces deux types de définitions ne représentent pourtant pas toujours bien la réalité. Tout d'abord, les prédictions respectives de formation de multivalents et de bivalents pour les auto- et les allopolyploïdes ne sont généralement valides que pour les polyploïdes récents. En effet, avec le temps, les polyploïdes tendent à « diploïdiser » leur génome et même les autopolyploïdes finissent par former des bivalents (Wendel 2000). Aussi, les premières définitions d'autopolyploïdes considéraient que ceux-ci évoluaient strictement par dédoublement de chromosomes (Stebbins 1980). Cependant, nous savons maintenant que ce mode de formation est peu répandu dans la nature. En fait, il semble que la plupart des polyploïdes soient formés par l'intermédiaire de gamètes non réduits (Ramsey et Schemske 1998 ; Burton et Husband 2000). Ainsi, même les autopolyploïdes pourraient provenir de l'hybridation d'individus génétiquement différenciés bien qu'à priori conspécifiques. Il existe donc plutôt un continuum de possibilités entre le doublement de génomes strictement identiques et celui de génomes hautement différenciés tout comme un continuum entre des patrons de ségrégation disomique et « complètement » polysomique. Grant (1981) a proposé la classification suivante pour mieux rendre compte de ces « continua ».

## A. Les autopolyploïdes qui contiennent :

- 1) Les autopolyploïdes au sens stricte AAAA (stricte duplication du génome à partir d'un individu)
- 2) Les autopolyploïdes dits « inter-raciaux » AAAA (croisement d'individus très proches)

## B. Les amphiploïdes qui contiennent :

- 1) les « *Segmental allopolyploid* » AsAsAtAt (issus du croisement d'individus différenciés mais dont la proximité ou la relative homologie des génomes permet la formation de multivalents)
- 2) les « *Genomic allopolyploid* » AABB (issus du croisement d'individus au génome très différenciés)
- 3) et les allo-autopolyploïdes: AAAABB

Aussi, des données moléculaires ont permis de démontrer que la plupart des taxons polyploïdes étaient formés à plusieurs reprises (Soltis et al. 2004) et avaient donc une origine multiple. En d'autres termes, on trouve souvent plusieurs cytotypes pour un même taxon polyploïde. Si de tels exemples provenaient initialement d'allopolyploïdes (Soltis et Soltis 1993, 1999), des travaux récents ont pu montrer que les autopolyploïdes ne dérogeaient pas à la règle (Soltis et al. 1989 ; Van Dijk et Bakx-Schotman 1997 ; Segraves et al. 1999 ; Joly et Bruneau 2004). Ainsi, il devient parfois difficile de détecter si on est face à un autopolyploïde avec une origine multiple ou résultant de l'hybridation de populations génétiquement différenciées mais conspécifiques ou très proches (allopolyploïde segmentale selon la classification de Grant). *Ipomoea batatas* appartient très probablement à ces cas intermédiaires.

Au départ de ce travail, deux scénarios s'affrontaient pour expliquer l'origine du génome hexaploïde de *I. batatas* :

- 1) Selon certains auteurs la patate douce aurait une origine hybride impliquant au moins deux génomes distincts (A et B). L'analyse des caractères morphologiques amena Austin (1988) à proposer que les deux espèces impliquées étaient probablement *I. triloba* (génome B) et *I. trifida* (génome A). *Ipomoea batatas* serait alors un auto-allopolyploïde (AAAABB).
- 2) D'autres auteurs suggèrent que *I. batatas* dériverait de populations de *I. trifida* diploïdes par autopolyploïdisations successives (Kobayashi 1984) et serait donc un taxon autopolyploïde (AAAAAA).

Les résultats obtenus ne fournissent pas de preuves directes d'une origine allopolyploïde (au sens génomique) de *I. batatas*, c'est-à-dire impliquant plusieurs génomes très différenciés et donc appartenant à des espèces différentes. En particulier, *I. triloba* ne semble pas avoir participé à la formation du génome de *I. batatas*, discréditant ainsi le scénario 1 ci-dessus. Des travaux antérieurs avaient déjà sérieusement mis en doute la participation de *I. triloba*, faute de données génétiques (Rajapakse et al. 2003) ou cytogénétiques (Srisuwan et al. 2006) claires, mais aussi parce que l'étude de la morphologie du pollen a clairement démontré l'éloignement phylogénétique de *I. triloba* (Srisuwan 2006). Il est toutefois important de souligner que les outils moléculaires utilisés dans cette thèse sont limités pour véritablement écarter le scénario allopolyploïde (ou auto-allopolyploïde). L'histoire évolutive d'un polyploïde, bien souvent « réticulée », pose de nombreux problèmes méthodologiques pour les reconstructions phylogénétiques (Fortune et al. 2007). Les marqueurs chloroplastiques sont souvent nécessaires. En effet, ils permettent d'identifier la participation de lignées dont le signal a pu être perdu par recombinaison, réarrangement génomique ou ségrégation polysomique au niveau du génome nucléaire. Toutefois, ils nous renseignent uniquement sur les contributeurs maternels. C'est bien dans le génome nucléaire qu'il faut donc chercher l'intégralité de l'information. Cependant, les marqueurs nucléaires utilisés dans ce travail sont peu appropriés. Les séquences « ITS » sont peu variables dans notre cas et donc peu informatives. De plus, nous n'avons pas pu cloner les différentes versions en présence dans le génome de la patate douce (6 au total) et accédons là encore à une information partielle. Aussi, certains travaux ont pu montrer que les ITS étaient soumis à des phénomènes d'évolution concertée qui homogénéisent les différentes copies et brouillent l'origine initiale des progéniteurs (Wendel et al. 1995). De leur côté les microsatellites nucléaires donnent accès à une similitude globale entre génomes et non pas une véritable image des relations phylogénétiques entre espèces (Emshwiller et al. 2009). De plus, le codage binaire des microsatellites pour des plantes avec des niveaux différents de ploïdie est susceptible d'introduire un biais dans les analyses statistiques. Ainsi, la proximité génétique pourrait être biaisée par le niveau de ploïdie. Seule la réalisation d'une phylogénie adaptée, mettant en évidence les différentes copies homologues comme potentiellement homéologues présentes chez *I. batatas* et dans d'autres espèces de la série permettra de résoudre véritablement cette question. Le défi méthodologique consistera à identifier différents gènes nucléaires présents en une seule copie dans le génome (c'est-à-dire qui n'appartiennent pas à des familles multi-géniques dupliquées), puis d'en isoler les différentes copies. D'autres séquences chloroplastiques pourront aussi être recherchées pour améliorer la résolution de la phylogénie chloroplastique.

En l'absence de preuves flagrantes d'allopolyploïdie, force est de nous tourner vers un scénario « autopolyploïde ». Nos résultats réaffirment la proximité génétique entre *I. batatas* et *I. trifida*, déjà attestée par l'ensemble des données établies auparavant. Toutefois, *I. trifida* ne peut être considéré comme l'ancêtre ou le progéniteur direct de la patate douce, comme on le trouve parfois écrit dans la littérature, mais représente un taxon de proche parenté. Les progéniteurs diploïdes de *I. batatas* n'ont pas encore été identifiés. Ils n'ont pas encore été collectés ou ont peut-être disparu, supplantés par les formes polyploïdes. En revanche, l'existence de plusieurs lignées chloroplastiques différenciées au sein de *I. batatas* démontre clairement l'implication d'au moins deux contributeurs maternels distincts dans la formation du génome de *I. batatas*. Deux scénarios (non exclusifs) peuvent alors être envisagés : i) *I. batatas* a été formé par l'hybridation de plusieurs lignées distinctes (conspécifiques ou proches) ; ii) *I. batatas* est un complexe autopolyploïde avec une origine multiple. La difficulté de trancher entre ces deux scénarios vient du fait qu'on s'intéresse ici à un taxon en partie

cultivé, pour lequel l'image de la distribution actuelle de la diversité a pu largement être modifiée par les activités humaines de domestication et de diffusion.

## II. QUELLES RELATIONS ENTRE I. TRIFIDA ET I. BATATAS ? DE LA DIFFICULTE D'IDENTIFIER ET DE DELIMITER DES ESPECES

*" Je viens juste de comparer entre elles des définitions de l'espèce [...], il est vraiment comique de voir à quel point peuvent être diverses les idées qu'ont en tête les naturalistes lorsqu'ils parlent de l'espèce ; chez certains, la ressemblance est tout, et la descendance de parents communs compte pour peu de choses ; chez d'autres, la ressemblance ne compte pratiquement pour rien, et la création est l'idée dominante ; pour d'autres encore, la descendance est la notion-clé ; chez certains, la stérilité est un test infaillible, tandis que chez d'autres, cela ne vaut pas un sou. Tout cela vient, je suppose, de ce que l'on essaie de définir l'indéfinissable ". Ainsi écrivait Darwin à J. D. Hooker le 24 décembre 1856.*

Qu'est-ce qu'une espèce ? Vaste question qui agite les biologistes depuis qu'existe la volonté de réaliser une classification des organismes vivants. Cependant, il n'y a toujours pas de consensus réel sur ce que les biologistes considèrent comme une espèce et sur ce qu'elle doit représenter. Certains s'intéressent à une définition « opérationnelle » ou « pratique » de l'espèce, définition qui doit servir à identifier une espèce sur le terrain. D'autres se questionnent davantage sur les processus biologiques impliqués dans la spéciation. Alors que certains pensent que les espèces devraient être définies en fonction du flux génique (Mayr 1942, 2000), d'autres définissent l'espèce en regard des mécanismes reproductifs (Paterson 1985), parfois en termes de compétition (Ghiselin 1975), ou encore de distinctions morphologiques et génétiques (Mallet 1995). Au sens morphologique, l'espèce est définie comme un groupe d'individus partageant des critères diagnostiques propres et suffisamment différents morphologiquement d'autres groupes semblables. Cette définition est souvent opérationnelle mais peut poser problème lorsqu'il y a des variations importantes entre individus au sein d'une même espèce. D'autre part, le concept d'espèce biologique, définition processuelle, fait référence à des groupes de populations naturelles interféconds mais reproductivement isolés d'autres groupes semblables (Mayr 1942, 2000). Toutefois, la formation d'une nouvelle espèce est un processus dynamique et le maintien de flux de gènes en est souvent une étape intermédiaire (Rieseberg et Willis 2007 ; Petit et Excoffier 2009). Selon la définition phylogénétique, une espèce doit à terme correspondre à un groupe phylogénétique résolu et homogène, c'est-à-dire monophylétique (Mishler, et Theriot 2000). Toutefois, l'hybridation ou encore le tri incomplet des lignées pendant le processus de spéciation peut conduire à des arbres phylogénétiques discordants (Nichols 2001 ; Degnan et Rosenberg 2009). Chez les végétaux, ces différentes définitions sont régulièrement mises à mal en relation notamment avec les phénomènes relativement courants d'hybridation (qui peut conduire à la fusion de deux espèces) et de polyploïdisation (qui peut conduire à la spéciation sympatrique) (Rieseberg et Willis, 2007 ; Paun et al. 2010). Aussi, la domestication interroge les concepts d'espèce chez les plantes. Alors que la domestication induit souvent des changements morphologiques intenses entre le sauvage et le cultivé, les flux de gènes sont rarement totalement interrompus avec l'ancêtre sauvage (Ellstrand et al. 1999 ; Desplanque et al. 1999 pour la betterave ; Chen et al. 2004 pour le riz ; Duputié et al. 2007 pour le manioc, par exemple). Aussi, la mise en culture et la dispersion des cultivars peut mettre en contact des populations auparavant isolées et permettre ainsi l'hybridation et l'établissement de flux de gènes entre la forme cultivée et d'autres espèces sauvages proches (van Heerwaarden et al. 2010 pour le maïs, ou encore Delplancke et al. 2011 pour l'amandier). Rien de plus complexe donc, que d'essayer de délimiter des espèces au sein d'un complexe sauvage/cultivé polyploïde, potentiellement d'origine hybride.

Initialement, les auteurs de l'hypothèse 2) ont identifié les formes intermédiaires de *Ipomoea* en tant que *I. trifida* et ils proposaient alors la notion de « complexe *I. trifida* » qui regroupait l'ensemble des formes diploïdes et polyploïdes de la série *Batatas*, se croisant avec la patate douce. Selon cette hypothèse, *I. batatas* faisait référence à un taxon cultivé qui dérivait de la domestication de formes hexaploïdes de *I. trifida*. De leur côté, les taxonomistes (et auteurs de l'hypothèse 1) reconnaissent *I. trifida* et *I. batatas* comme deux espèces bien distinctes. La forme des sépales et de la corolle sont des critères pertinents pour leur identification mais qui nécessitent souvent l'œil averti d'un spécialiste (Austin 1988). Suivant cette définition, les populations polyploïdes au statut controversé ont été déterminées comme des formes d' *I. batatas* (Bohac et al. 1993). De façon concordante, nos travaux de génétique montrent que *I. batatas* (6X) et *I. trifida* (2X) se distinguent d'un point de vue génétique et correspondent notamment à des lignées d'haplotypes différentes. Aussi, les formes intermédiaires polyploïdes possèdent les mêmes haplotypes que *I. batatas* (à l'exception d'un spécimen portant un haplotype unique). Il est donc raisonnable de les classer dans ce même taxon. Il nous semble alors pertinent de délimiter l'espèce *I. trifida* à un taxon sauvage diploïde et *I. batatas* à un complexe polyploïde regroupant à la fois des formes sauvages et cultivées. Il n'en reste pas moins que ces deux espèces sont proches. L'isolement reproducteur, en dépit de niveaux de ploïdie différents, n'est d'ailleurs pas total (Diaz et al. 1996). On est ici probablement face à un processus de spéciation récent et toujours en cours.

Il n'est pas rare que les taxons cultivés soient au départ délimités comme une espèce à part, incitant dès lors à rechercher les ancêtres sauvages au sein d'autres espèces proches, alors qu'en réalité des formes sauvages de cette même espèce peuvent aussi exister. Ces formes « intermédiaires » représentent pourtant des maillons essentiels pour la reconstruction de l'histoire de la domestication. On peut citer l'exemple du manioc, où la découverte de populations sauvages de *Manihot esculenta* ssp. *flabellifolia* a remis en cause l'hypothèse initiale de « compilo-espèces » qui stipulait que le manioc avait été domestiquée par hybridation de nombreuses espèces dont *Manihot aesculifolia*, une espèce mexicaine (Olsen 2004). C'est aussi le cas de *Oxalis tuberosa*, une racine andine octoploïde, pour laquelle la découverte de populations sauvages polyploïdes capables de faire des tubercules a largement permis d'affiner notre compréhension de son histoire évolutive (Emshwiller et Doyle 2002 ; Emshwiller 2006 ; Emshwiller et al. 2009). C'est donc sur le terrain que se trouve la clé de notre énigme. Une meilleure compréhension de l'origine de *I. batatas* passera nécessairement par la prospection et l'analyse écologique, morphologique et génétique des formes intermédiaires polyploïdes de *I. batatas*.

Sur le plan « épistémologique », notre modèle d'étude est riche d'enseignements car il laisse à penser ce qui se trouve aux interstices. Là où la science doit établir des classes pour se pencher sur le réel, le réel nous montre souvent qu'il y échappe. Les définitions, les classements guident nos investigations mais parfois aussi restreignent le champ des possibles : « la méthode qui consiste à définir et à tout déduire d'une définition peut convenir aux sciences de l'esprit, mais elle est contraire à l'esprit même des sciences expérimentales » disait Claude Bernard (1878). A vouloir mettre des mots sur les choses, on en oublie parfois de les regarder comme elles sont.

### III. DOMESTICATION UNIQUE OU MULTI-LOCALE ?

A l'heure actuelle, notre compréhension de la domestication en Amérique tropicale indique qu'il s'agit d'un phénomène spatialement diffus, qui a pu avoir lieu dans des régions écologiquement très différentes (Piperno 2011), même si une majorité de plantes serait issue des basses terres des forêts saisonnières tropicales. Certaines espèces ont été domestiquées à plusieurs reprises à partir de populations sauvages différenciées. D'autres ont une origine unique. Par la suite, certaines plantes ont diffusé relativement rapidement sur une grande échelle. Domestiqué il y a environ 8500 ans, le manioc est retrouvé au Panama il y a 7 600 ans et sur les côtes péruviennes il y a 7 500 ans (Piperno 2009 et 2011). Le maïs rejoint lui aussi rapidement l'Amérique du Sud depuis le Mexique (Ranere et al. 2009). Au contraire, d'autres plantes ont peu diffusé et sont initialement restées localisées (Clement et al. 2010). Il semblerait donc que chaque plante ait suivi une trajectoire évolutive propre, échappant à tout

modèle simple. Il doit aussi être souligné qu'une domestication multi-locale n'implique pas nécessairement que l'agriculture ait été « inventée » plusieurs fois de façon indépendante. La diffusion de plantes domestiquées a pu simplement « inspirer » les populations à s'approprier à leur tour leur flore locale (Piperno 2011).

Au sein des formes cultivées de *I. batatas*, on différencie sur la base des marqueurs nucléaires et chloroplastiques les variétés de la région Pérou-Equateur de celles de la région allant du Mexique à la Colombie. Un tel signal phylogéographique a pu être hérité des populations sauvages, impliquant dès lors deux domestications « indépendantes ». Des patrons très similaires ont pu être mis en évidence pour d'autres plantes, notamment au sein du genre *Phaseolus* (Chacon et al. 2005 ; Motta-Aldana et al. 2010). En effet, *Phaseolus vulgaris* et *Phaseolus lunatus* auraient été chacun domestiqués indépendamment en Amérique Centrale et dans les vallées andines, à partir de populations sauvages isolées et différenciées génétiquement. Cette différenciation des populations sauvages refléterait un phénomène de spéciation allopatrique en cours, à mettre en relation avec l'histoire tectonique de la région (fermeture de l'isthme de Panama et orogénèse alpine notamment) (Chacon et al. 2005 et 2007 ; Serrano-Serrano et al. 2010). Il est notable que les formes domestiquées dans ces différentes régions sont distinctes morphologiquement, forment des « races » et ne s'hybrident pas entre elles (Singh et al. 1991 ; Beebe et al. 2000 et 2001). L'histoire de la patate douce pourrait bien ressembler à celle des haricots. Dans ce cas, nos deux lignées haplotypiques correspondraient à deux populations différenciées et l'hypothèse d'une origine autoploïde multiple serait la plus probable. Cependant, nous n'observons pas de races de patate douce correspondant à ces deux groupes génétiques, tout comme ces populations ne semblent pas présenter de signes d'isolement reproducteur. On ne peut donc exclure dans notre cas que le signal phylogéographique observé soit plutôt le résultat de la diffusion de la plante par les hommes et que la patate douce ait plutôt une origine hybride, et soit issue d'une domestication unique.

Au sujet de l'origine de la patate douce, Dolores Piperno (2011) écrivait dans une revue sur la domestication en Amérique tropicale :

*“There is a major root crop for which neither Central American nor South American origins have been conclusively demonstrated.”*

A l'issue de cette thèse, la question de l'origine géographique de la domestication de la patate douce reste ouverte.

## IV. A LA RECHERCHE DE PREUVES GENETIQUES DE L'HYPOTHESE TRIPARTITE

Reconstruire l'histoire de la patate douce en Océanie est avant tout un travail « historiographique ». Il s'agit de mettre en ordre des faits, des événements passés et de reconstituer une intrigue particulière cohérente au travers de ces « champs événementiels ». L'hypothèse tripartite constitue le scénario le plus probable aujourd'hui sur l'introduction de la patate douce en Océanie, c'est-à-dire l'hypothèse qui maximise la cohérence des différents faits observables rassemblés. Jusqu'à aujourd'hui, les données assemblées ont été surtout d'ordre archéologique, historique, linguistique et ethnobotanique. Que nous apporte la génétique?

### IV.1. Les variétés polynésiennes et la voie Kumara

Grâce aux échantillons d'herbier, nous montrons que les variétés de patate douce présentes jusqu'au début du 20<sup>ème</sup> siècle en Polynésie ont clairement une signature génétique sud-américaine, c'est-à-dire qu'elles dérivent directement des variétés de la région Pérou-Equateur. Ainsi nos données génétiques corroborent les données linguistiques et apportent une preuve supplémentaire à l'existence d'au moins une connexion préhistorique entre la Polynésie et l'Amérique du Sud ayant permis le transfert de matériel végétal entre ces deux régions (la voie Kumara).

Les données historiques laissent penser que plusieurs variétés étaient déjà présentes à l'arrivée des Européens, au moins trois « lignées » distinctes (Green 2005). Aussi, des données linguistiques suggèrent l'existence de plusieurs transferts préhistoriques de plantes, c'est à dire de plusieurs contacts entre les Polynésiens et les populations d'Amérique. Il existe notamment une variété hawaïenne, globuleuse à chaire blanche, qui ne ressemble pas aux autres kumaras de Polynésie selon les dires des premiers explorateurs (Green 2005), et qui aurait pu être introduite indépendamment depuis les côtes colombiennes (Rensch 1991) ou encore celles d'Amérique centrale (Montenegro 2007). Malheureusement, nous n'avons pas pu identifier de planches d'herbier très anciennes à Hawaï. La réponse se trouve peut-être dans les nombreuses accessions d'herbier restantes au Bishop Museum de Honolulu (que nous n'avons pas eu le temps de génotyper dans cette thèse), collectées dans les années 1930 par Handy et qui pourraient bien contenir un représentant de cette variété. En l'état, nous ne pouvons donc conclure sur le nombre de contacts préhistoriques.

Aussi, la voie Kumara telle qu'elle est formulée dans l'hypothèse tripartite propose un itinéraire très détaillé du transfert de la patate douce en Polynésie, qui se serait réalisé relativement rapidement entre 1100 et 1300 AD. Nous ne disposons pas dans ce travail d'un échantillonnage suffisamment exhaustif pour permettre cette reconstruction fine. Il est peu probable que la génétique puisse d'ailleurs véritablement « pister » le transfert des lignées introduites, d'abord par manque de marqueurs polymorphes et de structuration de la diversité génétique à cette échelle de temps, mais aussi car la patate douce ne suit pas une évolution purement clonale. Les différentes variétés recombinaient entre elles une fois introduites et on perd ainsi rapidement leur signature génétique d'origine.

## IV.2. D' hier à aujourd' hui

Autre fait remarquable, nous montrons que les kumaras de Polynésie orientale jusqu'au début du 20<sup>ième</sup> siècle sont différentes (changement d'haplotypes) de celles qu'on trouve plus tard (dès les années 1960) et encore aujourd'hui dans cette même région. Il y aurait donc eu un remaniement de la base génétique initiale au 20<sup>ième</sup> siècle. Nous formulons ainsi l'hypothèse que des introductions plus tardives qui datent de l'aire européenne, ont recombinaient avec le pool local initial, modifiant ainsi la base génétique « neutre » du départ. En revanche, il est probable que les phénotypes reconnus par les cultivateurs et les noms associés (c'est-à-dire les déterminants « culturels » des variétés) aient été maintenus. Le même scénario a probablement eu lieu à Hawaï et en Nouvelle Zélande, mais faute d'un échantillon conséquent de spécimens d'herbier, nous ne pouvons le révéler. Pour tester véritablement cette hypothèse sur l'ensemble de la Polynésie, il faudrait pouvoir suivre en parallèle, le nom, la morphologie, et la génétique des kumaras dans le temps. Au début du 20<sup>ième</sup> siècle des agronomes comme Handy pour Hawaï, Best pour la Nouvelle Zélande ou encore Brown en Polynésie orientale ont recensé, collecté puis décrit les différentes variétés en présence tout comme réalisé des planches d'herbier (en tout cas Handy et Brown, non vérifié pour Best). Celles de Handy sont conservées au Bishop Museum à Hawaï et la description précise des variétés peut être trouvée dans son livre « *the Hawaiian Planters* » (1940). Le travail mené sur les herbiers dans cette thèse confirme la possibilité et l'importance de ce type d'échantillon dans la reconstruction de l'histoire de diffusion d'une plante cultivée, mais reste préliminaire, ouvrant la porte à de futures recherches.

## IV.3. A l' Ouest du Pacifique

Selon l'hypothèse tripartite, l'Ouest du Pacifique a reçu principalement des variétés de patate douce issues des voies historiques Batata et Camote (Yen 1974). Mais aussi dans certaines régions, notamment en Mélanésie orientale et en Nouvelle Guinée, des kumaras polynésiennes auraient pu être introduites. Toutefois, il n'existe aucune preuve archéologique à ce jour (Allen 2005). Les variétés d'Amérique tropicale du Mexique et des Caraïbes ne sont pas distinctes génétiquement, dès lors il devient impossible de différencier les variétés « Batata » des variétés « Camote » dans la zone d'introduction. En outre, le remaniement de la base génétique initiale au fil des nouvelles introductions

(mis en évidence à l'Est du Pacifique) a probablement aussi affecté l'Ouest du Pacifique, effaçant ainsi progressivement la signature génétique d'origine des premières introductions et notamment celles des kumaras s'il y en eu. Il est important de rappeler que les explorateurs espagnols Mendana et Quiros auraient pu introduire des variétés de type « kumara » (puisque provenant d'Amérique du Sud), très tôt dans le Pacifique (dès le 16<sup>ième</sup>), compliquant d'autant plus la résolution de l'origine de la patate douce dans cette région. L'image actuelle de la diversité montre que les variétés de l'Ouest du Pacifique ont une origine centraméricaine plutôt que sud-américaine et forment un groupe relativement homogène (même base génétique initiale), sans structure géographique intra-groupe apparente, comme attendu sous l'hypothèse tripartite. L'introduction de kumaras n'est pas montrée. Toutefois, nous n'avons pas pu réunir un échantillon de spécimens d'herbier représentatif pour cette région et donc nous ne pouvons pas écarter totalement l'idée d'introductions préhistoriques par les Polynésiens, « *l'absence de preuves n'étant pas la preuve de l'absence* ».

#### IV.4. D' Est en Ouest, d' Ouest en Est

Au regard des mouvements d'hommes et donc de matériel végétal qui ont eu lieu en Océanie, notamment depuis l'arrivée des Européens, on aurait pu s'attendre à une certaine homogénéisation de la diversité entre l'Ouest et l'Est du Pacifique, en dépit d'une origine différente. Pourtant on ne retrouve pas de variétés mélanésiennes en Polynésie et *vice versa*. Aussi, les populations de patate douce des deux régions restent relativement différenciées. Existe-t-il une frontière immatérielle entre ces deux mondes, frontière imperméable aux mouvements des végétaux ? Cette frontière existe d'un point de vue culturel (même si les Mélanésiens ne forment pas un groupe homogène, les Polynésiens eux forment une unité culturelle distincte). Les variétés étant des objets « sociaux », leur structuration génétique doit refléter l'organisation sociale et culturelle des sociétés qui les reproduisent (Leclerc et Coppens 2011). Dans notre cas, cette différenciation est aussi le reflet d'histoires d'introductions distinctes. Il serait intéressant d'analyser l'identité variétale des « enclaves polynésiennes », îles et atolls dont les habitants appartiennent culturellement au monde polynésien mais qui sont situées selon la géographie traditionnelle en Mélanésie, pour évaluer l'importance du facteur culturel dans l'organisation des échanges de végétaux et la structuration de la diversité génétique des plantes cultivées.

Qu'apporte donc la génétique ? Elle confirme tout bonnement les fondements de l'hypothèse tripartite. Elle s'ajoute aux autres types de preuves et vient consolider le scénario proposé par Barrau en 1975, développé par Yen en 1974 puis revu par Green en 2005. Certains auraient pu attendre de la preuve ADN qu'elles résolvent ou démontrent définitivement. La preuve ADN semble être pourvue d'un niveau de scientificité plus important que d'autres types de preuves. Elle serait en quelque sorte « la reine des preuves ». Pourquoi cette croyance ? Peut être car la preuve ADN requiert un niveau de technicité plus important et qu'aujourd'hui Science et Technique sont pratiquement devenues indissociables. Il est alors bon de méditer la phrase de De Candolle dans son ouvrage, l'origine des plantes cultivées (1882): « *Il est nécessaire d'user de différents moyens lorsqu'on veut s'assurer de l'origine des plantes cultivées. C'est, pour chaque espèce, une recherche dans le genre de celles que font les historiens et les archéologues, recherche variée, dans laquelle on se sert tantôt d'un procédé et tantôt d'un autre, pour les combiner ensuite et les apprécier selon leur valeur relative. Le naturaliste n'est plus ici dans son domaine ordinaire d'observations et de descriptions. Il doit s'appuyer sur des preuves testimoniales, dont il n'est jamais question dans les laboratoires* ».

De façon plus inquiétante, cette même tendance s'observe aussi au niveau sociétal, notamment dans le domaine de la criminalité et de la justice. La preuve ADN y est devenue la caution ultime de la « vérité judiciaire » et l'on oublie parfois de « s'encombrer » d'autres types de preuves. Aussi, le fichage ADN est aujourd'hui en plein essor. Alors que la découverte de l'ADN est l'une des plus belles « découvertes » de la biologie du XXI<sup>ème</sup> siècle, et que les applications de la génétique semblaient pouvoir contribuer au « bien être de l'humanité » (notamment au travers des applications thérapeutiques), la génétique se trouve ici mobilisée au service d'un système répressif plus qu'émancipateur, comme l'Histoire nous l'a déjà enseigné par le passé.



## V. CREER DE LA DIVERSITE CHEZ UNE PLANTE A PROPAGATION CLONALE

A partir de son étude sur la variation des caractères morphologiques des patates douces océaniques, Yen (1974) a très tôt suspecté que la reproduction sexuée a été la principale source de diversification variétale dans le Pacifique. Nos résultats génétiques confirment cette hypothèse et montrent que la plupart des variétés présentes aujourd'hui sont issues d'événements de recombinaison distincts et proviennent de la sélection locale de plantes issues de graines (chapitres 4 et 5). La patate douce rejoint ainsi bon nombre d'autres plantes à multiplication végétative qui elles aussi combinent reproductions sexuée et clonale.

### V.1. Fonctionnement(s) des systèmes mixtes

Dans ces systèmes mixtes, les agriculteurs tirent partie des deux systèmes de reproduction. D'un côté, la reproduction sexuée permet de purger les allèles délétères accumulés au fil du temps, mais aussi d'assembler au sein d'un même individu les mutations favorables survenues indépendamment dans différentes lignées, générant ainsi rapidement de nombreux variants potentiellement intéressants pour l'agriculteur (McKey et al. 2010). Elle permet donc à priori une évolution plus rapide que dans un système purement clonal. De son côté, la reproduction clonale permet en premier lieu de fixer ces variants d'intérêt sur un temps long. Mais elle donne aussi l'opportunité de sélectionner de nouveaux variants dérivés de mutations somatiques ou de modifications épigénétiques héréditaires et contribue elle aussi à générer de la diversité variétale. Il est probable que la diversité créée par ces deux systèmes n'ait pas la même « fonction » pour le paysan. La reproduction sexuée pourrait permettre de sélectionner plus rapidement des variants pour des traits « adaptatifs », qui ont un déterminisme génétique complexe (souvent multi-génique) comme la résistance aux maladies ou encore l'adaptation aux conditions abiotiques du milieu (sécheresse, froid...). La diversification clonale, elle, aurait plutôt un rôle dans la création de variants pour des traits « d'usage » ou traits « culturels », comme la couleur, le goût, la texture ou encore la forme de l'organe consommé ou du reste de la plante. Il est important de souligner que la reproduction sexuée produit certes des génotypes de bonne qualité, mais conduit aussi à la formation de génotypes de mauvaise qualité (dépression de consanguinité, perte de l'hétérozygotie, accumulation de mutations délétères...) en cassant les combinaisons alléliques favorables (McKey et al. 2010). Le rôle des paysans dans la sélection et l'intégration des plantules est donc crucial pour évincer ces variants « de mauvaise qualité ».

L'étude de ces systèmes mixtes pour différentes plantes, notamment le manioc en Amazonie (revue dans McKey et al. 2012 a et b) et en Afrique (Delêtre 2008), l'igname d'Afrique de l'Ouest (Scarcelli 2006), ou encore le taro du Vanouatou (Caillon 2006), a mis en évidence des fonctionnements différents, où la reproduction sexuée prend plus ou moins d'importance. Chez le manioc, les plantes issues de graines sont trouvées fortuitement lors de la mise en place d'un nouveau champ et elles correspondent à des hybrides au sein du compartiment cultivé (hybrides inter-variétaux ou issus de croisement de plantes appartenant à la même variété). Au terme d'un long processus sélectif (cf. introduction pour les modalités), les plantules les plus hétérozygotes sont fréquemment intégrées aux variétés proches morphologiquement, formant ainsi des groupes variétaux polyclonaux. Plus rarement, elles sont conservées pour créer de nouvelles variétés. Dans le cas du manioc, c'est donc la diversité intra-variétale qui se trouve enrichie par ce processus d'intégration des plantules. Chez l'igname, les plantes issues de graines peuvent provenir d'un croisement entre variétés mais souvent aussi d'un croisement entre formes cultivées et formes sauvages, permettant ainsi l'introduction de nouveaux allèles dans le système. Les plantules sont en général gardées pour créer de nouvelles variétés, enrichissant ainsi la diversité variétale. Cette intégration est un événement relativement rare comparé au cas de figure typifié par le manioc. Aussi, la sélection de la plantule se fait en forêt (milieu naturel) et non plus au champ (milieu cultivé). Chez le taro, les plantules sont rares, naissent essentiellement d'hybridations inter-variétales, sont utilisées pour générer de nouvelles variétés et comme pour le manioc sont trouvées lors de l'ouverture d'un nouveau bassin laissé en

jachère de nombreuses années. Un point commun à ces différents systèmes mixtes semble tout de même se dégager. L'intégration des plantules et donc le maintien et l'utilisation de la reproduction sexuée pour une plante à multiplication végétative ne sont pas gérée de façon « volontaire » par le paysan. Les cultivateurs ne croisent pas des génotypes pour créer de nouvelles variétés mais tirent profit de dynamiques « naturelles » et ce plus ou moins consciemment. Il faut aussi souligner qu'au sein même d'une espèce, la gestion de ces systèmes mixtes peut aussi varier selon la localité. Quels sont donc les facteurs qui influencent les modalités de coexistence de la reproduction sexuée et clonale dans les agro-systèmes?

## **V.2. Les facteurs qui influencent le fonctionnement des systèmes mixtes**

### ***V.2.1. La biologie et l'écologie de la reproduction de la plante***

Un des premiers facteurs qui influe sur les modalités de coexistence entre reproduction sexuée et clonale a trait aux caractéristiques intrinsèques de la biologie et de l'écologie de la reproduction de la plante. Si chez le manioc la probabilité d'incorporer des recombinants est élevée, c'est d'abord le reflet d'une reproduction sexuée particulièrement active comparativement à d'autres plantes à racines ou tubercules. Souvent, les plants de manioc ont amplement le temps de fleurir et de produire des fruits et des graines avant d'être récoltés puisqu'ils sont stockés sur pied et donc laissés longtemps en terre après maturité. Ce mode de stockage en champ est aussi valable pour notre modèle d'étude, la patate douce. En revanche, les pieds d'ignames ou encore de taros sont entièrement arrachés lors de la récolte, limitant dès lors le temps nécessaire à la floraison. Les conditions environnementales peuvent aussi s'avérer décisives pour l'induction de la floraison chez certaines espèces. Dans le cas de la patate douce, il est clair que la température (froid) et la durée du jour (jours courts) favorisent la floraison et donc que certains environnements comme les hautes terres de Nouvelle Guinée seront plus favorables que d'autres à la recombinaison. Par ailleurs, les opportunités de croisement ne sont pas identiques pour toutes les plantes. L'igname est une espèce dioïque qui nécessite la présence de pieds mâles et femelles pour produire des graines. De même, la patate douce est auto-incompatible et la production de graines est dépendante de la coexistence de variétés possédant des groupes d'auto-incompatibilité différents. Il est probable que les premières introductions préhistoriques de patate douce en provenance d'Amérique du Sud aient introduit initialement un nombre restreint de génotypes et de groupes d'incompatibilité. Des introductions supplémentaires ont alors été nécessaires pour véritablement mettre en marche le cycle de diversification par voie sexuée. En outre, les fruits du manioc contiennent de grosses graines qui tombent à proximité des pieds, attendant des conditions favorables à leur germination (près du sol) suite à la ré-ouverture du milieu cultivé. C'est aussi le cas de la patate douce, dont les graines sont relativement grandes. A l'inverse, les graines de taro sont très petites. Elles sont en général dispersées dans les forêts secondaires environnantes (et non pas en milieu cultivé), ce qui rend moins probable la survie de la plantule (problèmes de compétition) tout comme leur détection par les paysans. Pour certaines espèces, la coexistence de formes avec des niveaux de ploïdie différents peut aussi présenter un frein à la reproduction sexuée. Ainsi des « biologies » différentes vont laisser plus ou moins de place au maintien de la reproduction sexuée dans le système clonal.

### ***V.2.2. Les caractéristiques « génomiques »***

Les mutations somatiques (ou les modifications épigénétiques) sont en général considérées comme des événements assez rares et donc l'évolution purement clonale comme relativement lente. Pour obtenir rapidement de nouvelles variétés, les paysans auraient alors intérêt à se tourner vers la sélection de plantules. La fréquence des mutations dépend pourtant de l'espèce considérée. Notamment, les génomes polyploïdes « récents » sont hautement dynamiques et sujets à d'importants réarrangements génomiques (Adams et Wendel 2005). Ils doivent donc présenter un fort taux de mutation et l'évolution clonale serait alors efficace. A ce titre, la patate douce est connue pour avoir un taux de

mutation somatique particulièrement élevé (Lewthwaite 2004). Aussi, les croisements impliquant des génomes polyploïdes produisent de nombreuses combinaisons alléliques (notamment si la ségrégation est polysomique) et créent ainsi des descendance avec de nombreux variants potentiels. Pour ces plantes, même si les événements de reproduction qui donnent naissance à une plantule viable sont rares, la descendance est potentiellement très diversifiée morphologiquement, ce qui attirerait efficacement l'œil du cultivateur pour la sélection de nouvelles variétés. C'est ce qui se passe probablement pour la patate douce. Au contraire, dans le cas du manioc, il y a, semble-t-il, beaucoup de plantules produites à chaque cycle. En revanche, une partie présentera peu de différences avec les variétés « parents », ce qui pourrait d'ailleurs contribuer à expliquer pourquoi les paysans les classent dans des variétés existantes plutôt que de les utiliser pour créer de nouvelles variétés.

### *V.2.3. Les conditions de l'environnement*

Dans un contexte environnemental instable comme celui des îles du Pacifique, sujets notamment à de nombreuses catastrophes naturelles (sécheresses, cyclones...) qui peuvent détruire l'intégralité des récoltes, les graines - notamment celles dans la banque de graines du sol - peuvent permettre de pallier à un éventuel manque de matériel de propagation (Caillon 2006). L'utilisation de la reproduction sexuée peut être particulièrement importante à la suite à de tels événements. Aussi, dans les milieux hétérogènes, comme les montagnes des hautes terres de Nouvelle Guinée par exemple, la reproduction sexuée permettrait d'obtenir rapidement des variants adaptés aux conditions locales, ce qui prendrait plus de temps dans un système purement clonal.

### *V.2.4. Les pratiques agricoles et la relation culturelle à la « diversité »*

On peut distinguer deux types de facteurs « humains » qui influencent l'importance et les modalités de l'intégration de la reproduction sexuée dans le système de propagation clonale.

Tout d'abord, il y a les pratiques culturelles que nous qualifierons de facilitatrices, c'est à dire celles qui vont accroître la fertilité sexuelle et augmenter le nombre de plantules produites. Parmi elles, on trouve notamment le mode de plantation. En effet pour ces espèces allogames, l'obtention de graines est d'autant plus fréquente que les cultures sont poly-variétales. Le temps de croissance laissé à la plante est aussi un facteur déterminant. Les systèmes culturels à longue révolution qui font place à des jachères sur plusieurs années et conservent les anciens plants favorisent l'accomplissement du cycle complet de reproduction sexuée.

Enfin, la diversification par voie sexuée dépend des modalités de sélection et d'intégration de ces plantules, c'est-à-dire du choix exercé par le cultivateur. L'intérêt des cultivateurs pour les plantes issues de graines relève d'une approche véritablement expérimentale (Mc Key et al. 2012). En effet, une jeune plantule n'est pas agronomiquement intéressante dans un premier temps. C'est seulement après plusieurs cycles de multiplication clonale que le paysan peut évaluer l'intérêt de la plantule. Il semble que l'intégration de ces plantules relève dans de nombreux cas plus des perceptions et des représentations paysannes de la diversité cultivée que d'une démarche motivée par des fins agronomiques. Ce choix reflète finalement la relation « culturelle » du paysan à la diversité, à l'expérimentation et à l'innovation. Caillon et Lanouguère-Bruneau (2005) ont montré qu'à Vanua Lava, une petite île au Nord du Vanouatou, les différents clones de taros sont de véritables supports de mémoire, mémoire des ancêtres, des lieux. La capture de sauvageons représente pour ces populations une occasion nouvelle d'écrire leur histoire et est donc socialement valorisée. Au contraire, lors de précédents travaux menés sur l'île de Tanna (une île au Sud du Vanouatou), j'ai pu observer des comportements différents à l'égard de plantules entre des plantes « traditionnelles » (ignames et taros) et celles plus récemment introduites qui n'ont pas acquis un statut culturel élevé (manioc et patate douce). A Tanna, les variétés de taro ou d'ignames ainsi que la fertilité des récoltes associées sont placés sous l'auspice d'un « magicien » et contribuent à un patrimoine biomagique bien délimité que chaque village ou groupe de villages conserve (Bonnemaison 1997). Il se pourrait que ce système magique contraigne les possibilités de création variétale et notamment d'expérimentations

individuelles, en encadrant de façon très claire le stock de variétés à reproduire. Pour les plantes libérées de cette contrainte magique, l'innovation et l'expérimentation pourraient être facilitées.

Dans le cas de la patate douce, il semble que l'on ait un système plus clonal en Polynésie qu'en Mélanésie. Nous avons en tout cas détecté plus de clones et de lignées clonales à l'Est du Pacifique. On ne peut exclure que cet effet soit lié aux stratégies d'échantillonnage des collections *ex situ* utilisées dans ce travail et ne reflète pas réellement une différence en terme de gestion paysanne. Toutefois, ce même constat peut être dressé pour bon nombre d'autres plantes à propagation clonale en Océanie (cf. introduction). Les Polynésiens gèreraient ainsi des systèmes plus strictement clonaux que les Mélanésiens. Nous avons montré en introduction que pour certaines plantes amenées de l'Ouest par les premiers migrants, on observait un gradient positif de stérilité d'Ouest en Est, gradient hérité du processus même de colonisation d'île en île. Nous tenons là une explication « biologique » au phénomène observé. On aurait ainsi plus de clonalité à l'Est car les cultivateurs gèreraient des plantes moins fertiles à cause d'une stérilité génétique ou écologique (cf. introduction). On peut en parallèle regarder cette différence Est-Ouest de gestion des reproductions sexuée et clonale d'un point de vue « culturel ». Il est possible que les Mélanésiens et les Polynésiens entretiennent des relations différentes au « sauvageon » (peut être parce qu'ils les rencontrent moins souvent). Les plantules, produits de la reproduction sexuée, sont en effet inscrites dans une dynamique naturelle de régénération. Elles appartiennent au « sauvage ». La propagation clonale, elle, est bien le fait de l'homme. Ainsi cette différence Est-Ouest traduit des modes différents d'appropriation de l'environnement, des rapports de « domesticité » distincts entre ces deux régions. L'adaptation des sociétés océaniques aux îles lointaines de Polynésie est associée à une intensification de l'aménagement du territoire par l'Homme et à une « artificialisation » plus importante des lieux (Yen 1990). Ainsi, les polynésiens pourraient laisser moins de place à ce qui relève du « spontané », du « naturel », que ne le font les Mélanésiens. Ils seraient figés « culturellement » dans un système plus clonal en dépit des possibilités offertes par la reproduction sexuée. Ce n'est qu'une hypothèse de travail qu'il serait intéressant de tester à partir de l'étude d'autres plantes introduites plus récemment, comme le manioc par exemple. Bien sûr, une explication ne prévaut pas sur l'autre et l'état de la diversité d'une plante cultivée se trouve justement être un compromis évolutif à l'interaction de ces déterminants biologiques et culturels.

## **VI. RACONTER L'HISTOIRE DES HOMMES EN OCEANIE AU TRAVERS DE CELLES DES ESPECES COMMENSALES OU DOMESTIQUES : DES PISTES DE REFLEXION**

La question de l'origine des peuples Océaniques et des voyages préhistoriques qui leur permirent de s'installer dans ces myriades d'îles éparpillées dans l'Océan Pacifique taraudent marins et scientifiques depuis les premières découvertes. Initialement, ce sont les travaux d'archéologie menés depuis plus d'une centaine d'années qui ont permis de lever le voile sur l'origine des hommes du Pacifique, reconstruisant pas à pas la chronologie des migrations. La linguistique, notamment historique, dont l'objet est de rendre compte de l'évolution et de la généalogie des langues, a souvent permis de préciser l'origine des migrants, profitant aussi des méthodes quantitatives de reconstruction cladistique issues de la biologie. Plus récemment, la génétique humaine s'est elle aussi penchée sur cette histoire. L'analyse comparée de marqueurs génétiques variés (notamment des marqueurs mitochondriaux à hérédité maternelle, des marqueurs situés sur la région non recombinante du chromosome Y (NRY) à hérédité paternelle, et des marqueurs autosomaux avec une double hérédité) a souvent confirmé les données issues des autres disciplines, mais aussi dévoilé des patrons inattendus et contradictoires, révélant des échanges de gènes entre groupes, parfois insoupçonnés (cf. introduction). L'entreprise menée dans cette thèse s'intéresse non plus aux hommes directement mais tente de reconstruire cette histoire au travers de celle des espèces domestiques ou commensales qui ont accompagnés les hommes dans leurs parcours migratoires. Depuis longtemps les patrons de

distribution des espèces cultivées ont servi à inférer la mobilité des hommes en Océanie (Barrau et al. 1961), mais l'avènement des méthodes moléculaires a conduit à revitaliser ce champ de recherche.

La force de ce type d'approche réside en premier lieu dans le fait qu'il peut nous renseigner sur des mouvements ou interactions entre populations n'ayant pas ou peu impliqué d'échanges de gènes entre groupes humains, donc difficilement détectables par la génétique humaine (Storey et al. 2011). C'est notamment le cas des échanges présumés entre les Polynésiens et les populations d'Amérique dont il est question dans ce travail. Il existe d'autres plantes qui pourraient s'avérer intéressantes dans l'étude de tels contacts, comme la tomate polynésienne ou encore le savonnier, tous deux d'origine américaine mais présentes anciennement en Polynésie (cf. introduction). Aussi, un faisceau grandissant de preuves laisse penser qu'il y a eu de nombreuses interactions durant l'Holocène entre les îles d'Asie du Sud-est, l'Océanie proche et la Micronésie (Donohue et Denham 2010). Une hypothèse récente stipule même que l'origine des Polynésiens est à rechercher dans le contact entre les Lapitas, installés en Polynésie occidentale depuis, et des peuples venus d'Asie du Sud-est, passés par la Micronésie et ayant débarqués au Samoa/Tonga il y a environ 2000-1500 ans (Addison et Matisoo-Smith 2010). Ces interactions ont pu être démontrées grâce à la reconstruction génétique et linguistique de l'histoire évolutive de la banane (Perrier et al. 2011). Mais, il reste là encore de nombreuses espèces qui pourraient être utilisées pour caractériser ces interactions, citons entre autres le taro, l'igname, la canne à sucre, l'arbre à pain ou encore le murier à papier. Aussi, elles ont plusieurs avantages techniques ou méthodologiques comparés notamment à la génétique humaine et s'avèrent particulièrement utiles lorsque les restes archéologiques sont rares (Storey et al. 2011). Travailler sur de l'ADN humain pose en général un certain nombre de problèmes de contamination, en particulier pour l'analyse de restes archéologiques. En outre, l'accès au matériel archéologique est parfois contrôlé par les communautés locales (Storey et al. 2011). Il peut ainsi être plus aisé de constituer un échantillonnage représentatif d'une zone pour les espèces domestiques que pour les humains qui les peuplent. Enfin, au-delà de la reconstruction des routes migratoires, ces approches permettent d'analyser les modalités de diffusion et d'interactions culturelles dans cette région et nous questionnent sur les « choix humains » en contexte migratoire. Pourquoi avoir ramené principalement de la patate douce d'Amérique et non pas du manioc, ou encore du maïs ? Il semblerait que la ressemblance entre la patate douce et l'igname aurait permis une intégration « rapide » de cette nouvelle plante, qui aurait pu être traitée d'un point de vue culturel mais aussi agricole comme un « igname métaphorique » (Leach 2005). Ces approches offrent ainsi une porte d'entrée sur l'évolution et l'adaptation culturelle des sociétés océaniques.

Le travail développé dans cette thèse confirme la pertinence de ce type d'approche et permet en même temps d'en souligner certaines limites.

## **VI.1. La nécessité d'une approche temporelle**

On suspecte depuis longtemps l'influence de l'aire européenne sur les patrons initiaux de distribution des espèces cultivées et commensales. Des études menées sur le rat ont permis de mettre en évidence de tels changements temporels (Matisoo-Smith et Robins 2009). De façon similaire, notre étude démontre qu'il y a eu un remaniement de la base génétique initiale de la patate douce au cours du temps. Il apparaît donc plus que nécessaire d'associer des échantillons anciens pour ce type d'études et de constituer un échantillonnage qui permette un suivi spatio-temporel de la diversité. Les spécimens d'herbier représentent à ce titre un matériel tout à fait intéressant et notamment ceux des herbiers historiques. Pour certaines plantes, il existe aussi des restes archéobotaniques potentiellement exploitables, comme pour la gourde par exemple (revu par Clarke 2009). Enfin, d'autres sources de matériel peuvent être envisagées. Il y a par exemple dans le cas du murier à papier, la possibilité d'extraire de l'ADN à partir des étoffes anciennes ou « Tapa », conservés dans de nombreux musées dans le monde (Seelenfreud et al. 2010).

## **VI.2. Le difficile choix des marqueurs moléculaires**

Le choix des marqueurs est particulièrement complexe pour ce type d'étude, car il doit correspondre à une délicate adéquation entre niveau de polymorphisme et représentativité de l'échantillonnage. En effet, l'histoire des interactions et échanges qu'on souhaite révéler dans le Pacifique est relativement récente et il est donc dans certains cas difficile de trouver des marqueurs suffisamment polymorphes à cette échelle temporelle. Dans notre cas, les marqueurs utilisés (notamment les marqueurs chloroplastiques) ne permettent pas véritablement de pister les clones introduits par les polynésiens puis distribués dans le triangle polynésien et donc de reconstruire en détail l'histoire de cette diffusion. Aussi, travailler avec de l'ADN ancien restreint le choix des marqueurs. Il faut des marqueurs courts (< 150pb) et spécifiques de l'espèce qu'on étudie pour limiter les problèmes de contamination. A ce titre, les microsatellites représentent un compromis intéressant qui doit être résolutif dans de nombreux cas. Pour les espèces polyploïdes, il y a toutefois un « vide statistique » sur le traitement des données microsatellites nucléaires qui en limite la résolution. A l'avenir, les approches de séquençage de nouvelle génération devraient apporter des outils moléculaires plus efficaces.

## **VI.3. La nécessité d'une approche transversale**

Notre travail réaffirme la nécessité d'une approche transversale et pluridisciplinaire pour aborder ces questions et étudier des objets aux confins du biologique et du culturel. Nous montrons que les différentes composantes d'une variété, c'est-à-dire ses caractères phénotypiques, son nom et son génotype ne suivent pas nécessairement la même trajectoire évolutive. Alors que les bases génétiques neutres peuvent être remaniées rapidement (évolution rapide), les déterminants culturels (le nom et le phénotype), c'est-à-dire ceux sur lesquels l'homme agit directement, qu'il peut sélectionner, peuvent être conservés (évolution plus lente). Ceci illustre que transmission (ou diffusion) génétique et culturelle suivent des modalités et rythmes d'évolution distinctes, que les « *memes* » (pour reprendre l'expression de Richard Dawkins (1976) ou « *semes* » (Hewlett et al. 2002) d'un côté, et les gènes, de l'autre, ne se transmettent pas de façon identique. Seule une approche pluridisciplinaire permet donc de réconcilier ces différentes visions d'une même histoire.

# **VII. QUELLES IMPLICATIONS POUR LA GESTION DES RESSOURCES GENETIQUES?**

## **VII.1. La conservation et l'amélioration des plantes à racines et tubercules**

Consciente des risques d'érosion de la biodiversité cultivée, la communauté scientifique internationale s'est mobilisée dès les années 1960s-1970s pour inventorier et collecter le maximum de ressources génétiques pour les espèces présentant un intérêt économique majeur. La patate douce a bénéficié de ce statut et plus de 35 000 accessions de patates douces ont été assemblées et sont maintenues dans différentes collections *ex situ* de par le monde (FAO 2011). La méthode de conservation la plus répandue pour ces plantes dites à graines récalcitrantes est la conservation aux champs. Cette approche exige de grands espaces de culture, des moyens financiers importants et du personnel « qualifié » pour maintenir un étiquetage cohérent des variétés sur le long terme. Il n'est pas toujours aisé de réunir toutes ces conditions dès lors qu'on s'écarte des grandes institutions de recherche et de conservation internationales (principalement le CIP dans le cas de la patate douce). Aussi, ces collections *ex situ* ont pour vocation première d'être des réservoirs de gènes pour l'amélioration variétale et nécessitent donc d'être évaluées et caractérisées, augmentant encore les coûts associés. Ces collections offrent déjà plus d'accessions que les améliorateurs ne sont en mesure d'évaluer et les moyens actuellement alloués

sont disparates et insuffisants (FAO 2011). Il y a donc à l'heure actuelle un réel besoin de diminuer le nombre d'accessions en collection et d'établir des collections dites représentatives (Frankel et Brown 1984), c'est-à-dire de définir un échantillon limité d'accessions issues d'une collection plus vaste, choisies pour représenter au mieux le spectre de diversité existant.

Les programmes d'amélioration variétale attenants à ces collections *ex situ* sont souvent peu accessibles aux agriculteurs des zones marginales (c'est-à-dire aux conditions environnementales particulières dans lesquelles se pratique principalement une agriculture de subsistance), zones où se concentrent notamment la culture vivrière de la patate douce. En effet, les produits de l'amélioration variétale élaborés en station, souvent homogènes, sont peu adaptés aux conditions de culture locales, qui elles sont hétérogènes. De même, les variétés améliorées ne sont pas en mesure de répondre à l'ensemble des exigences culturelles et sociales, à la fois utilitaires mais aussi identitaires, de ces sociétés traditionnelles et sont souvent laissées pour compte par les agriculteurs. C'est par exemple le cas en Papouasie-Nouvelle-Guinée, où les agriculteurs se refusent à utiliser des variétés améliorées de patate douce, préférant leurs cultivars sélectionnés localement (Van Wijmeersch 2001). En outre, pour les plantes à propagation végétative comme la patate douce, l'amélioration variétale est relativement complexe d'un point de vue scientifique (Lebot 2002, 2009). On manipule (croise) des plantes souvent très hétérozygotes, ce qui se traduit par une forte variabilité des caractères dans la descendance d'hybrides, ainsi que la ré-apparition de caractères sauvages délétères et oblige dès lors à cribler un très grand nombre de descendants. Pour un polyploïde, la sélection récurrente est d'autant plus longue à cause du masquage potentiel des allèles (Soltis et Soltis 2000). Aussi, pour la plupart de ces plantes, l'obtention d'hybrides peut être rendue relativement complexe par une sexualité qu'il est parfois difficile de maîtriser en station.

Ces différents constats ont mené depuis le début des années 1980s à la mise en œuvre de méthodes de conservation dites « à la ferme ». Les approches réunies sous ce terme revêtent en réalité une diversité de formes (Wood et Lenné 1997, FAO 2011). Les systèmes traditionnels ont pu être envisagés comme relativement conservateurs statiques, notamment dans le cas des plantes à multiplication végétative, garant de la conservation d'une diversité « héritée », de la reproduction de variétés figées. Cette vision peut conduire à des méthodes de conservation à la ferme qui « mettent sous cloche » ces systèmes traditionnels. Plusieurs études de cas ont par la suite permis de mettre en évidence que dans ces systèmes clonaux, la diversité maintenue n'était en réalité pas statique, mais bel et bien dynamique et évolutive (Elias et al. 2000 ; Caillon 2006 ; Scarcelli 2006 ; Sardos 2008) . Plus que de conservation à la ferme, il s'agirait alors de mettre en place des méthodes de gestion et même d'amélioration participative qui conserve le potentiel adaptatif de ces systèmes et fortifie leurs adaptabilités et leurs résiliences (Lebot et al. 2005). Une méthode de Distribution Géographique de la Diversité Allelique (DGDA) a été proposée dans ce sens et vise à conserver de la diversité sous forme d'allèles plutôt que sous la forme de génotypes fixés (Lebot et al. 2005). Cette approche se décompose en quatre étapes : 1) la composition d'un échantillon représentatif de la diversité utile de l'espèce, 2) la distribution géographique des génotypes en vue d'élargir les bases génétiques locales, 3) la distribution d'allèles sous forme de clones (ou de graines) issus de descendance en ségrégation et 4) la sélection locale par les producteurs de variants localement adaptés. Cette méthodologie générale est fondée sur une complémentarité entre conservation *ex situ* (étape 1, 2 et 3) et à la ferme (étape 4). Sa mise en place nécessite une bonne connaissance à la fois de la structure de la diversité à une échelle globale et des processus de diversification à l'œuvre dans les agro-systèmes en question. Les résultats présentés dans ce travail permettent quelques recommandations pour la mise en œuvre des approches de conservations *ex situ* et à la ferme pour la patate douce.

## VII.2. Quelle diversité « utile » pour cette espèce ?

La réalisation de collections représentatives pose la délicate question de ce qui doit être conservé : des allèles, des génotypes particuliers, de la diversité fonctionnelle, un potentiel adaptatif ? En effet, les critères pour l'élaboration de telles collections sont difficiles à « hiérarchiser ». Surtout, il n'est pas évident d'identifier aujourd'hui ce qui se révélera « utile » demain. La méthode classique consiste à stratifier la collection selon plusieurs critères pertinents (eco-géographiques, agro-morphologiques notamment) et à échantillonner le maximum de diversité à l'intérieur des groupes (Frankel et Brown

1984). En général, les caractères agro-morphologiques utiles sont prioritaires mais ce sont aussi ceux pour lesquelles l'évaluation est la plus lourde. Les marqueurs moléculaires neutres sont aujourd'hui largement utilisés pour la constitution de ces collections (Bataillon et al. 1996). Ils permettent en premier lieu d'éliminer les doublons. Aussi, les patrons de structure génétique qu'ils révèlent peuvent servir de base pour la stratification des collections.

Au terme de cette étude, il semble judicieux de considérer la diversité cultivée de notre espèce (dans la zone d'étude bien sûr) autour de quatre pôles. Dans la zone d'origine, les groupes Nord et Sud représentent deux entités génétiques (sur la base des deux types de marqueurs) et géographiques distinctes et doivent potentiellement contenir des gènes d'intérêt distincts. En Océanie, deux groupes sont à considérer, les variétés de l'Est du Pacifique (triangle polynésien) et le groupe de l'Ouest (Mélanésie, Papouasie-Nouvelle-Guinée et Asie du Sud-Est) qui ont des histoires évolutives distinctes et sont restés relativement isolés. Les marqueurs moléculaires permettent de proposer un premier niveau de stratification que les données morphologiques ne laissaient pourtant pas transparaître. En choisissant des individus particulièrement hétérozygotes, le nombre de clones à conserver pour représenter le spectre de diversité allélique de chaque groupe géographique peut être limité.

A l'intérieur de chaque groupe, nos marqueurs neutres semblent toutefois peu résolutoires pour accéder à une structuration génétique à une échelle plus fine. Pourtant, certaines variétés, même au sein d'un groupe génétique, présentent des possibilités agronomiques différentes et semblent localement adaptées à des conditions agro-écologiques particulières. Notamment, les variétés des hautes-terres de Nouvelle Guinée ne sont pas productives dans les conditions de basse-terre (Van Wijmeersch 2001). Aussi, il existe des disparités importantes en termes de résistance à certaines maladies comme la galle causée par le champignon *Elsinoe batata*, en relation avec la virulence des souches fongiques (Lenné et al. 1994). La sélection paysanne a pu produire au fil du temps des variétés « locales », qu'il paraît essentiel d'identifier et de conserver, cette fois-ci en tant que génotype à l'intérieur de chaque groupe. Plus que d'allèles utiles, il est probable que pour notre espèce polyploïde ce soit des combinaisons alléliques particulières générées par la recombinaison génétique, qui soient « utiles ». Dans les génomes polyploïdes particulièrement, la recombinaison est susceptible d'induire des interactions géniques particulières ou encore des effets de dosage des différents allèles en présence et de générer rapidement de nombreux variants, élargissant ainsi le « potentiel adaptatif » de la plante. Leur criblage nécessite des méthodes « classiques » d'évaluation agro-morphologique. Cette idée d'adaptation locale doit toutefois être nuancée. Il est souvent admis que les variétés traditionnelles sont localement adaptées en relation avec la sélection exercée par les cultivateurs et l'environnement. Toutefois, cela a été rarement démontré (Wood et Lenné 1997). La sélection paysanne suit aussi d'autres critères, notamment organoleptiques, qui sont parfois prépondérants et peuvent conduire à une perte de variation pour les caractères « adaptatifs ». Aussi, les variations des conditions de l'environnement en milieu « traditionnel » peuvent conduire à ce que les paysans sélectionnent plutôt des variétés avec un large spectre adaptatif que des variants localement adaptés. Chez la patate douce, la polyploïdie est certainement l'une des raisons du succès adaptatif de la plante à une large gamme d'environnements en regard d'autres cultures.

Ce manque de structuration génétique à une échelle régionale est le reflet du processus de diversification sous-jacent pour notre espèce. L'exemple du Pacifique montre qu'à partir d'un nombre relativement restreint d'introductions, la plante a pu s'adapter en un temps relativement court (environ 300 ans pour les hautes terres de Papouasie-Nouvelle-Guinée) à des conditions très variées de culture et produire une multitude de formes différentes de par l'usage de la recombinaison génétique par les cultivateurs. Conserver de la diversité utile pour l'avenir, c'est donc donner la possibilité à ce haut potentiel évolutif et adaptatif de se maintenir. C'est s'assurer la possibilité d'une reproduction sexuée active. Pour ce faire, il est important de se focaliser sur la conservation de génotypes hétérozygotes et florifères et de favoriser aussi des génotypes qui possèdent des groupes d'auto-incompatibilité différents et peuvent ainsi s'hybrider.

Dans cette discussion sur la diversité « utile » de l'espèce *I. batatas*, nous nous sommes focalisés sur la diversité « utile » au sens de l'améliorateur, c'est-à-dire la diversité fonctionnelle qui garantit la résilience des agro-écosystèmes et le potentiel adaptatif de l'espèce. Il ne faut pourtant pas négliger les autres dimensions de cette diversité qu'on peut qualifier de socio-culturelles. Il y a d'abord celle qui est liée aux usages de la plante (diversité organoleptique par exemple) mais aussi celle qui reflète l'identité sociale et culturelle des sociétés traditionnelles. Les menaces sur la



biodiversité ne sont pas seulement d'ordre environnemental, mais tout autant sociétal. La déstructuration des sociétés traditionnelles conduit inmanquablement à une érosion de la diversité cultivée: reproduire des plantes permet de reproduire des sociétés qui à leur tour peuvent reproduire des plantes. Ce qui est « utile » est donc aussi ce qui a du sens socialement et culturellement et peut être reconnu et reproduit par les cultivateurs. Nous n'avons pas abordée véritablement cette question au cours de cette thèse mais insistons sur l'importance de considérer cette dimension de la diversité pour la mise en œuvre de sa conservation.

Enfin, lorsqu'on parle de diversité « utile » pour une espèce cultivée, il est aussi nécessaire d'envisager les formes sauvages qui peuvent être mobilisées pour la recherche de nouveaux gènes d'intérêt (Spillane et Gepts 2001). De nombreuses espèces sauvages de *Ipomoea* apparentées à la patate douce sont maintenues dans des collections *ex situ* sous forme de graines (plus de 1100 accessions maintenues au CIP notamment). Pourtant, les croisements entre *I. batatas* et la plupart de ces espèces sont peu efficaces et l'introgession de gènes d'intérêt depuis le pool sauvage limitée. Nous montrons dans cette thèse qu'il doit exister des populations sauvages polyploïdes de *I. batatas* en Amérique tropicale, formes qui devraient être prioritaires en terme de conservation de la diversité « utile » pour cette espèce.

### **VII.3. De la possibilité de mettre en œuvre des méthodes de gestion à la ferme et d'amélioration participative pour la patate douce en Océanie**

L'histoire de la patate douce en Océanie démontre l'existence d'une forte dynamique de création variétale où les différentes introductions ont pu être intégrées au fil du temps et la diversité locale remaniée, recombinaison au grès de ces introductions. Nous avons identifié très peu d'accessions d'Amérique tropicale dans le Pacifique, tout comme nous n'avons pas retrouvé aujourd'hui de génotypes identiques aux échantillons d'herbier du début du 20<sup>ème</sup> siècle. Il est donc probable qu'il y ait un fort renouvellement des variétés (ou en tout cas des génotypes) au cours du temps. Face à cette dynamique, la conservation à la ferme ne semble pas pouvoir avoir comme objectif le maintien de variétés fixées. A ce titre, la conservation *ex situ* doit être mise en place de façon complémentaire pour atteindre cet objectif. Par contre, l'image des processus passés met en évidence l'existence d'un potentiel « biologique » dans ces agro-systèmes pour une gestion et une amélioration à la ferme, telles que proposées dans la DGDA. En effet, des clones judicieusement choisis peuvent être distribués, puis devraient être recombinés localement « à la ferme ». Se pose tout de même la question de l'échelle temporelle. Notre étude révèle l'efficacité et la dynamique de ces systèmes à une échelle de plusieurs centaines d'années, trois cents ans par exemple pour les hautes terres de Papouasie-Nouvelle-Guinée. Nous avons toutefois très peu d'idées à l'issue de ce travail sur la rapidité, ni sur les rythmes même de ces processus. La diversification variétale est-elle un processus continu dans le temps, ou est-elle hétérogène, comme cela semble être le cas à une toute autre échelle, celle de la diversification des espèces (Gould et Eldredge 1972) ? Les changements que connaît la région Pacifique actuellement sont relativement rapides (augmentation de la densité de population, changements climatiques, nouvelles pressions pathogènes,...cf. introduction) et donc le développement et l'intégration de cultivars « adaptés » à ces nouvelles conditions doivent l'être aussi (Siméoni et Lebot 2012). Il semble donc important de trouver les moyens « d'accélérer », de renforcer le processus de diversification. L'améliorateur peut d'abord participer aux choix de génotypes qui peuvent augmenter la fréquence des événements de recombinaison dans le système clonal (génotypes florifères, groupes d'autoincompatibilité différents, maximisation de la diversité allélique injectée...). Aussi, il peut contribuer à mieux comprendre et valoriser les pratiques « facilitatrices » (champs polyvariétaux, conservation sur pied, ...). En revanche, la réussite de ce type de méthodologie dépend *in fine* du choix des cultivateurs d'intégrer ou non des plantules à leurs stocks de variétés locales. Il est important de garder à l'esprit que les cultivateurs ne sont pas des améliorateurs au sens « scientifique » du terme, c'est-à-dire qu'ils ne réalisent pas volontairement des croisements entre génotypes dans l'idée d'un progrès génétique selon des objectifs ciblés. Ils exploitent une dynamique spontanée et il est probable que leurs pratiques relèvent plus d'un attrait pour l'expérimentation et l'innovation que d'une démarche d'amélioration. Les expériences de

conservation *in situ* rapportent qu'il est souvent difficile de faire prendre conscience aux cultivateurs de l'intérêt de la diversité au sens de l'améliorateur, pour la résilience de leurs agro-systèmes. Par contre, il pourrait être utile d'aborder cette question de la diversité sous l'angle de l'expérimentation. Une compréhension fine des processus socio-culturels qui construisent le rapport des cultivateurs à la diversité semble une étape préalable tout à fait nécessaire pour la mise en place de ce type de méthodologie.

#### VII.4. Ressources génétiques et appropriation du vivant

Alors que se mettait en place au 19<sup>ième</sup> siècle le régime des brevets et des protections des inventions dans différents domaines, le vivant semblait rester tacitement exclu du champ de la propriété intellectuelle. Sa gratuité, tout comme sa capacité infinie de reproduction et d'évolution, empêchaient semble-t-il d'en concevoir l'appropriation. Le vivant était une chose commune, une « *res communis* » (Azam 2008). Conformément à la séparation entre le monde animé et le monde inanimé, généralement admise dans la culture occidentale, les brevets s'appliquaient au monde inanimé et ne pouvaient concerner le monde animé. Dans les années 1980s, l'avènement des biotechnologies va bouleverser ce paradigme. Le vivant est alors perçu comme de la matière biologique, un assemblage de pièces détachées (organes, cellules, gènes, molécules) que l'homme peut modifier à sa guise. Aussi, avec les progrès du génie moléculaire, le vivant devient un véritable enjeu économique et la diversité biologique (notamment la biodiversité cultivée) acquiert le statut de ressources génétiques, que l'on peut valoriser d'un point de vue économique. Dès lors, le vivant peut et doit s'approprier. Cette appropriation peut porter sur le procédé d'obtention, sur l'organisme lui-même dans son entièreté, ou bien encore sur certains de ses composants seulement, ses gènes par exemple. Les variétés végétales n'ont pas échappé à ce processus et c'est même dans ce domaine d'application que le concept de brevetabilité du vivant fait sa première apparition dans la législation. En 1930 aux Etats-Unis, le *Plant Patent Act* autorise déjà le brevetage de plantes obtenues « artificiellement » (par sélection artificielle) en supprimant la différence traditionnelle entre le vivant et l'inanimé et en la remplaçant par l'opposition entre le naturel et l'artificiel.

C'est dans cet esprit utilitariste et marchand que va être négociée la convention sur la diversité biologique signée à Rio en 1992 sur la conservation de la biodiversité. Elle a été conçue pour que tous les signataires partagent à la fois les efforts mais aussi les bénéfices de la conservation. La convention va notamment distribuer des droits entre les pays du Nord et du Sud, prémices nécessaires à la création d'un marché des ressources génétiques. À cette fin, elle propose d'abandonner la notion de patrimoine mondial de l'humanité au profit de la reconnaissance de la souveraineté des États (article 3 de la convention). Elle définit des droits d'usage et d'accès des ressources, où les droits de propriété intellectuelle sur le vivant deviennent des outils de conservation: d'une part, les droits de propriété intellectuelle (brevets) des industries du vivant sont reconnus (article 16.5), et d'autre part, les droits (à définir) des populations locales et indigènes sur leurs ressources génétiques et leurs savoirs doivent eux aussi être reconnus (article 8j). Chacun des usagers doit pouvoir s'approprier sa part de vivant.

Les variétés végétales n'échappent pas à la règle. Deux systèmes de propriété intellectuelle existent à l'heure actuelle pour la protection des variétés « artificielles », d'une part le brevet, mais aussi le COV (en Europe), pour Certificat d'Obtention Végétale. Le COV, contrairement au brevet, permet une utilisation libre des variétés (sans rémunération à l'obteneur) pour la création de nouvelles variétés. Pour qu'une variété soit enregistrée, plusieurs critères doivent être satisfaits : la distinction (se différencie par un caractère important des autres variétés), l'homogénéité (homogène pour l'ensemble des caractères) et la stabilité (identique dans le temps). Ce type de protection paraît tout à fait inadapté pour des variétés traditionnelles, qui justement sont hétérogènes et évoluent au cours du temps, mais semblent plutôt destinés à protéger les intérêts des obtenteurs des variétés végétales dans les pays industrialisés.

Pourtant, cette mouvance de reconnaissances des droits de propriété touche aussi les communautés locales qui veulent à leur tour accéder à une protection de leurs ressources. A ce titre, les communautés Maoris de Nouvelle Zélande revendiquent actuellement un droit de propriété intellectuelle sur leurs variétés de kumaras (trois variétés au total) au travers d'une requête plus globale de droits de propriété sur les savoirs traditionnels liés à la faune et la flore locales, adressée en

2006 au tribunal de Waitangi (créé en 1975 pour s'occuper en particulier des problèmes de droits des Maoris). La difficulté de ces objets particuliers, les variétés traditionnelles, est qu'ils sont à la frontière entre le biologique et le culturel. Les données obtenues dans ce travail illustrent bien cette complexité. Nous montrons que les kumaras Maoris ne sont probablement plus, d'un point de vue génétique, les mêmes que celles introduites initialement par les ancêtres Polynésiens. D'un point de vue biologique, elles seraient des « hybrides » entre les anciennes *kumaras* et des introductions européennes plus tardives. Peuvent-elles donc être la propriété des Maoris uniquement ? Pour autant, ces variétés sont sans conteste des « productions » culturelles Maoris, des marqueurs de l'identité Maori. Quel système de protection peut donc leur convenir ?

Face à ces objets complexes, il est toujours possible d'élargir le champ d'application de la propriété intellectuelle, d'élaborer de nouvelles formes de brevets et de protection. Ces objets complexes ne devraient-ils pas plutôt nous inciter à mettre fin à cette escalade d'appropriation du vivant et à imaginer d'autres moyens de partager, de mettre en commun, les ressources de notre environnement, en rupture avec les logiques marchandes ?

## *Références*

- Adams K.L. & Wendel J.F. (2005) Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology*, 8, 135-141
- Allaby R.G. (2007) Origins of plant exploitation in Near Oceania. A review. In: *Population genetics, linguistics and culture history in the south west Pacific* (ed. Friedlaender JS), pp. 181 - 198. Oxford University Press, Oxford
- Allaby R.G., Fuller D.Q. & Brown T.a. (2008) The genetic expectations of a protracted model for the origins of domesticated crops. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 13982-6
- Allen M.G. (2005) The evidence for sweet potato in islands Melanesia. In: *The Sweet Potato in Oceania: a Reappraisal* (eds. Ballard C, Brown P, Bourke RM & Hardwood T), pp. 99-108. Oceania Publications, Sydney
- Altieri M.A. (1999) The ecological role of biodiversity in agroecosystems. *Agriculture Ecosystems & Environment*, 74, 19-31
- Ames M. & Spooner D.M. (2008) DNA from herbarium specimens settles a controversy about origins of the European potato. *American Journal of Botany*, 95, 252-257
- Ammerman A.J. & Cavalli-Sforza L.L. (1984) *The Neolithic transition and the genetics of populations in Europe*. Princeton University Press, Princeton.
- Antheaume B. & Lawrence R. (1985) To Help or Too Much Help - Changing Food Crop Economies in the South-Pacific. *Etudes Rurales*, 367-387
- Arnaud-Haond S., Duarte C.M., Alberto F. & Serrao E.A. (2007) Standardizing methods to address clonality in population studies. *Molecular Ecology*, 16, 5115-5139
- Arrigo N., Felber F., Parisod C., Buerki S., Alvarez N., David J. & Guadagnuolo R. (2010) Origin and expansion of the allotetraploid *Aegilops geniculata*, a wild relative of wheat. *The New phytologist*, 187, 1170-1180
- Arrigo N., Guadagnuolo R., Lappe S., Pasche S., Parisod C. & Felber F. (2011) Gene flow between wheat and wild relatives: empirical evidence from *Aegilops geniculata*, *Ae. neglecta* and *Ae. triuncialis*. *Evolutionary Applications*, 4, 685-695
- Asfaw A., Blair M.W. & Almekinders C. (2009) Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces from the East African highlands. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 120, 1-12
- Austin D.F. (1977) Hybrid Polyploids in *Ipomoea* Section Batatas. *Journal of Heredity*, 68, 259-260
- Austin D.F. (1978) *Ipomoea*-Batatas Complex .1. Taxonomy. *Bulletin of the Torrey Botanical Club*, 105, 114-129
- Austin D.F. (1988) The taxonomy, evolution and genetic diversity of sweet potatoes and related wild species. In: *Exploration, maintenance, and utilization of sweet potato genetic resources - Report of the First Sweet Potato Planning Conference*, pp. 27-58. International Potato Center (CIP), Lima, Peru
- Austin D.F. (1991) *Ipomoea littoralis* (Convolvulaceae)-Taxonomy, Distribution, and Ethnobotany. *Economic Botany*, 45, 251-256
- Austin D.F. & Huaman Z. (1996) A synopsis of *Ipomoea* (Convolvulaceae) in the Americas. *Taxon*, 45, 3-38
- Azam G. « Les droits de propriété sur le vivant ». *Développement durable et territoires [En ligne]*, Dossier 10: Biens communs et propriété
- Badr A., Muller K., Schafer-Pregl R., El Rabey H., Effgen S., Ibrahim H.H., Pozzi C., Rohde W. & Salamini F. (2000) On the origin and domestication history of barley (*Hordeum vulgare*). *Molecular Biology and Evolution*, 17, 499-510
- Baert A. (2001) *Histoire de la découverte des régions australes: les Salomon, Marquises, Santa Cruz, Tuamotu, Cook du Nord et Vanuatu* l'Harmattan, Paris.
- Ballard C. (2005) Still good to think with: the sweet potato in Oceania. In: *The Sweet Potato in Oceania: a reappraisal*, pp. 5 - 13. Oceania Publications, Sydney
- Bandelt H.J., Forster P. & Röhl A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16, 37-48.
- Bardy K.E., Albach D.C., Schneeweiss G.M., Fischer M.a. & Schönswetter P. (2010) Disentangling phylogeography, polyploid evolution and taxonomy of a woodland herb (*Veronica*

- chamaedrys* group, Plantaginaceae s.l.) in southeastern Europe. *Molecular phylogenetics and evolution*, 57, 771-86
- Barnaud A., Deu M., Garine E., McKey D. & Joly H.I. (2007) Local genetic diversity of sorghum in a village in northern Cameroon: structure and dynamics of landraces. *Theoretical and Applied Genetics*, 114, 237-248
- Barrau J. L'énigme de la patate douce en Océanie. *Etudes d'Outre Mer*, 40, 3-4
- Barrau J. (1956) Le milieu et l'agriculture traditionnelle en Mélanésie. In: *Annales de Géographie*, pp. 362-382
- Barrau J. (1962) Les plantes alimentaires de l'Océanie : origines, distribution et usages. In: *Faculté des Sciences de Marseille*, Marseille
- Barrau J. (1965) "L'Humide et le sec": an essay on ethnobotanical adaptation to contrastive environments in the Indo-Pacific area. *Journal of the Polynesian Society*, 74, 329-346
- Barrau J. (1978) Domesticamento. In: *Enciclopedia Einaudi*, pp. 49 - 71. Einaudi, Torino
- Barton N.H. & Charlesworth B. (1998) Why sex and recombination? *Science*, 281, 1986-90
- Bataillon T.M., David J.L. & Schoen D.J. (1996) Neutral genetic markers and conservation genetics: Simulated germplasm collections. *Genetics*, 144, 409-417
- Baudouin L. & Lebrun P. (2009) Coconut (*Cocos nucifera* L.) DNA studies support the hypothesis of an ancient Austronesian migration from Southeast Asia to America. *Genetic Resources and Crop Evolution*, 56, 257-262
- Beebe S., Skroch P.W., Tohme J., Duque M.C., Pedraza F. & Nienhuis J. (2000) Structure of genetic diversity among common bean landraces of Middle American origin based on correspondence analysis of RAPD. *Crop Science*, 40, 264-273
- Beebe S., Rengifo J., Gaitan E., Duque M.C. & Tohme J. (2001) Diversity and origin of Andean landraces of common bean. *Crop Science*, 41, 854-862
- Bellwood P. (2012) Holocene Population History in the Pacific Region as a Model for Worldwide Food Producer Dispersals. *Current Anthropology*, 52, S363-S378
- Bohac J.R., Austin D.F. & Jones A. (1993) Discovery of wild tetraploid sweetpotatoes. *Economic Botany*, 47, 193-201
- Bonin A., Ehrich D. & Manel S. (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology*, 16, 3737-3758
- Bonnemaison J. (1996) *Les fondements géographiques d'une identité. L'archipel du Vanuatu. Essai de géographie culturelle. Livre II : Les gens des lieux. Histoire et géosymboles d'une société enracinée : Tanna*. Orstom, Paris.
- Boster J.S. (1984) Classification, cultivation and selection of Aguaruna cultivars of *Manihot esculenta* (Euphorbiaceae). In: *Ethnobotany in the Neotropics* (eds. Ghilleen T, Kallunki P & Kallunki JA), pp. 34-47. New York Botanical Garden, New York
- Bourke M. (2005a) The continuing Ipomoean revolution in Papua New Guinea. In: *The Sweet Potato in Oceania: a reappraisal*, pp. 171 - 180. Oceania Publications, Sydney
- Bourke M. & Studies A. (2009) Sweet Potato in Papua New Guinea. In: *The Sweetpotato*. Springer
- Bourke R.M. (2005b) Sweet Potato in Papua New Guinea: the plant and people. In: *The Sweet Potato in Oceania: a reappraisal*, pp. 15 - 24. Oceania Publications, Sydney
- Brand D.D. (1971) The sweet potatoe: an exercise in methodology. In: *Man accross the sea: problems of Pre-Colombian contacts* (eds. Riley CL, Kelley JC, Pennington CW & Rands RL), pp. 343 - 365. University of Texas Press, Austin
- Brown A.H.D. (1989) Core Collections - a Practical Approach to Genetic-Resources Management. *Genome*, 31, 818-824
- Brown F.B.H. (1935) *Flora of Southeastern Polynesia, volume III*. Bernice P. Bishop Museum, Honolulu.
- Brown P. & Brookfield H. (2005) Sweet potato, pigs and the Chimbu of the Papua New Guinea highlands. In: *The Sweet Potato in Oceania: a reappraisal*, pp. 131 - 136. Oceania Publications, Sydney
- Brush S., Kesseli R., Ortega R., Cisneros P., Zimmerer K. & Quiros C. (1995) Potato Diversity in the Andean Center of Crop Domestication. *Conservation Biology*, 9, 1189-1198

- Bryan G.J., McNicoll J., Ramsay G., Meyer R.C. & De Jong W.S. (1999) Polymorphic simple sequence repeat markers in chloroplast genomes of Solanaceous plants. *Theoretical and Applied Genetics*, 99, 859-867
- Bulmer R. (1965) Beliefs concerning the propagation of new varieties of sweet potato in two New Guinea Highland societies. *Journal of the Polynesian Society*, 74, 237-239
- Burger J.C., Chapman M.A. & Burke J.M. (2008) Molecular insights into the evolution of crop plants. *American Journal of Botany*, 95, 113-122
- Burle M.L., Fonseca J.R., Kami J.A. & Gepts P. (2010) Microsatellite diversity and genetic structure among common bean (*Phaseolus vulgaris* L.) landraces in Brazil, a secondary center of diversity. *Theoretical and Applied Genetics*, 121, 801-813
- Burnier J., Buerki S., Arrigo N., K  pfer P. & Alvarez N. (2009) Genetic structure and evolution of Alpine polyploid complexes: *Ranunculus kuepferi* (Ranunculaceae) as a case study. *Molecular Ecology*, 18, 3730-3744
- Burton T.L. & Husband B.C. (2000) Fitness differences among diploids, tetraploids, and their triploid progeny in *Chamerion angustifolium*: Mechanisms of inviability and implications for polyploid evolution. *Evolution*, 54, 1182-1191
- Buteler M.I., Jarret R.L. & LaBonte D.R. (1999) Sequence characterization of microsatellites in diploid and polyploid *Ipomoea*. *Theoretical and Applied Genetics*, 99, 123-132
- Caicedo A.L., Williamson S.H., Hernandez R.D., Boyko A., Fledel-Alon A., York T.L., Polato N.R., Olsen K.M., Nielsen R., McCouch S.R., Bustamante C.D. & Purugganan M.D. (2007) Genome-wide patterns of nucleotide polymorphism in domesticated rice. *PLoS Genet*, 3, 1745-56
- Caillon S., Quero-Garcia J., Lescure J.P. & Lebot V. (2006) Nature of taro (*Colocasia esculenta* (L.) Schott) genetic diversity prevalent in a Pacific Ocean island, Vanua Lava, Vanuatu. *Genetic Resources and Crop Evolution*, 53, 1273-1289
- Caillon S. & Degeorges P. (2007) Biodiversity: negotiating the border between nature and culture. *Biodiversity and Conservation*, 16, 2919-2931
- Casas A., Otero-Arnaiz A., P  rez-Negr  n E. & Valiente-Banuet A. (2007) *In situ* management and domestication of plants in Mesoamerica. *Annals of Botany*, 5, 1005-1101
- Cauvin J. (1997) *naissance des divinit  s, naissance de l'agriculture*. Paris CNRS Edition, Paris.
- Chacon M.I., Pickersgill B. & Debouck D.G. (2005) Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. *Theoretical and applied genetics*, 110, 432-44
- Chen L.J., Lee D.S., Song Z.P., Suh H.S. & Lu B.R. (2004) Gene flow from cultivated rice (*Oryza sativa*) to its weedy and wild relatives. *Annals of Botany*, 93, 67-73
- Chen Z.J. (2010) Molecular mechanisms of polyploidy and hybrid vigor. *Trends in Plant Science*, 15, 57-71
- Childe V.G. (1942) *What happened in history*, Harmondsworth.
- Clark L.V. & Jasieniuk M. (2011) POLYSAT : an R package for polyploid microsatellite analysis. *Molecular Ecology Resources*, 11, 562-566
- Clarke A.C., Burtenshaw M.K., McLenachan P.A., Erickson D.L. & Penny D. (2006) Reconstructing the origins and dispersal of the polynesian bottle gourd (*Lagenaria siceraria*). *Molecular Biology and Evolution*, 23, 893-900
- Clarke A.C. (2009) Origins and Dispersal of the Sweet Potato and Bottle Gourd in Oceania: Implications for Prehistoric Human Mobility. In: Massey University, Palmerston North, New Zealand
- Clement C.R., de Cristo-Ara  jo M., D'Eeckenbrugge G.C., Alves Pereira A. & Pican  o-Rodrigues D. (2010) Origin and domestication of native Amazonian crops. *Diversity*, 2, 72-106
- Clement M., Posada D. & Crandall K.A. (2000) TCS : a computer program to estimate gene genealogies. *Molecular Ecology*, 9, 1657-1659
- Cohen M.N. (1977) *The Food Crisis in Prehistory: Overpopulation and the Origins of Agriculture*. Yale University Press, New Haven.
- Coil J. & Kirch P. (2005) An Ipomoean landscape: archeology and the sweet potato in Kahikinui, Maui, Hawaiian Islands. In: *The Sweet Potato in Oceania: a reappraisal*, pp. 71-84. Oceania Publications, Sydney

- Comai L. (2005) The advantages and disadvantages of being polyploid. *Nature Reviews Genetics*, 6, 836-46
- Coppend d'Eeckengrugge G. & Duval I.g.p. (2009) The domestication of pineapple : context and hypotheses. *Pineapple news*, 16, 15 - 27
- Cui L., Wall P.K., Leebens-Mack J.H., Lindsay B.G., Soltis D.E., Doyle J.J., Soltis P.S., Carlson J.E., Arumuganathan K., Barakat A., Albert V.A., Ma H. & dePamphilis C.W. (2006) Widespread genome duplications throughout the history of flowering plants. *Genome Res*, 16, 738-49
- Darwin C. (1968) *The variation of plants and animals under domestication* John Murray, London.
- Dawkins R. (1976) *The Selfish Gene*. Oxford University Press.
- De Candolle A. (1883) *Origine des plantes cultivées*. Germer Baillière, Paris.
- De Langhe E. & De Maret P. (1999) Tracking the banana : its significance in Early Agriculture. In: *The Prehistory of Food* (eds. Godsen C & Hather J), pp. 377-396, Routledge
- De Silva H.N., Hall A.J., Rikkerink E., McNeilage M.A. & Fraser L.G. (2005) Estimation of allele frequencies in polyploids under certain patterns of inheritance. *Heredity*, 95, 327-334
- Degnan J.H. & Rosenberg N.A. (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, 24, 332-340
- Delêtre M. (2010) The ins and outs of cassava diversity in Gabon, Central Africa: a pluridisciplinary approach to the dynamics of genetic diversity of *Manihot esculenta* Crantz (Euphorbiaceae). In: *Trinity College*. University of Dublin
- Delplancke M., Alvarez N., Espindola A., Joly H., Benoit L., Brouck E. & Arrigo N. (2012) Gene flow among wild and domesticated almond species: insights from chloroplast and nuclear markers. *Evolutionary Applications*, 5, 317-329
- Denham T. (2007) Early to Mid-Holocene plant exploitation in New Guinea : Towards a contingent interpretation of agriculture. In: *Rethinking agriculture - Archaeological and ethnoarchaeological perspectives* (eds. Denham T, Iriarte J & Vrydaghs L), pp. 78 - 108. Left Coast Press, Walnut Creel, Cal.
- Denham T. (2011) Early Agriculture and Plant Domestication in New Guinea and Island Southeast Asia. *Current Anthropology*, 52, S379-S395
- Denham T.P., Haberle S.G., Lentfer C., Fullagar R., Field J., Therin M., Porch N. & Winsborough B. (2003) Origins of agriculture at Kuk Swamp in the Highlands of New Guinea. *Science*, 301, 189-193
- Desplanque B., Boudry P., Broomberg K., Saumitou-Laprade P., Cuguen J. & Van Dijk H. (1999) Genetic diversity and gene flow between wild, cultivated and weedy forms of *Beta vulgaris* L- (Chenopodiaceae), assessed by RFLP and microsatellite markers. *Theoretical and Applied Genetics*, 98, 1194-1201
- Diamond J. (1997) Location, location, location: The first farmers. *Science*, 278, 1243-1244
- Diamond J. (2002) Evolution, consequences and future of plant and animal domestication. *Nature*, 418, 700-707
- Diamond J. (2003) Guns, Germs, and Steel in 2003. *Antipode*, 35, 829-831
- Diamond J. & Bellwood P. (2003) Farmers and their languages: The first expansions. *Science*, 300, 597-603
- Diaz J., Delapiente F. & Austin D.F. (1992) Enlargement of Fibrous Roots in *Ipomoea* Section Batatas (Convolvulaceae). *Economic Botany*, 46, 322-329
- Diaz J., Schmiediche P. & Austin D.F. (1996) Polygon of crossability between eleven species of *Ipomoea*: Section Batatas (convolvulaceae). *Euphytica*, 88, 189-200
- Dickau R., Ranere A.J. & Cooke R.G. (2007) Starch grain evidence for the preceramic dispersals of maize and root crops into tropical dry and humid forests of Panama. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 3651-3656
- Dobney K. & Larson G. (2006) Genetics and animal domestication: new windows on an elusive process. *Journal of Zoology*, 269
- Doebley J., Stec A. & Hubbard L. (1997) The evolution of apical dominance in maize. *Nature*, 386, 485-8
- Doebley J.F., M. M. Goodman, and C. W. Stuber. (1984) Isoenzymatic variation in *Zea* (Gramineae). *Systematic Botany*, 9, 203-218

- Donohue M. & Denham T. (2010) Farming and Language in Island Southeast Asia Reframing Austronesian History. *Current Anthropology*, 51, 223-256
- Doust A.N., Devos K.M., Gadberry M.D., Gale M.D. & Kellogg E.A. (2004) Genetic control of branching in foxtail millet. *Proceedings of the National Academy of Science U S A*, 101, 9045-50
- Doust A.N. (2007) Grass architecture: genetic and environmental control of branching. *Current Opinion in Plant Biology*, 10, 21-5
- Doyle J.J., Flagel L.E., Paterson A.H., Rapp R.A., Soltis D.E., Soltis P.S. & Wendel J.F. (2008) Evolutionary genetics of genome merger and doubling in plants. *Annual Reviews of Genetics*, 42, 443-61
- Dubcovsky J. & Dvorak J. (2007) Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science*, 316, 1862-6
- Duputie A., David P., Debain C. & McKey D. (2007) Natural hybridization between a clonally propagated crop, cassava (*Manihot esculenta* Crantz) and a wild relative in French Guiana. *Molecular Ecology*, 16, 3025-3038
- Duputié A., Massol F., David P., Haxaire C. & McKey D. (2009) traditional Amerindian cultivators combine directional and ideotypic selection for sustainable management of cassava genetic diversity. *Journal of Evolutionary Biology*, 22, 1317-1325
- Eldredge N. & Gould S.J. (1972) Punctuated Equilibria: An Alternative to Phyletic Gradualism. In: *Models in Paleobiology* (ed. Schopf T, J.M.), pp. 82-115. Freeman, Cooper and Co., San Francisco
- Elias M., Panaud O. & Robert T. (2000a) Assessment of genetic variability in a traditional cassava (*Manihot esculenta* Crantz) farming system, using AFLP markers. *Heredity*, 85, 219-230
- Elias M., Panaud O. & Robert T. (2000b) Assessment of genetic variability in a traditional cassava (*Manihot esculenta* Crantz) farming system, using AFLP markers. *Heredity (Edinb)*, 85 Pt 3, 219-30
- Elias M., Penet L., Vindry P., McKey D., Panaud O. & Robert T. (2001) Unmanaged sexual reproduction and the dynamics of genetic diversity of a vegetatively propagated crop plant, cassava (*Manihot esculenta* Crantz), in a traditional farming system. *Molecular Ecology*, 10, 1895-1907
- Elias M., Muhlen G.S., McKey D., Roa A.C. & Tohme J. (2004) Genetic diversity of traditional south american landraces of cassava (*Manihot esculenta* Crantz): An analysis using microsatellites. *Economic Botany*, 58, 242-256
- Ellstrand N.C., Prentice H.C. & Hancock J.F. (1999) Gene flow and introgression from domesticated plants into their wild relatives. *Annual Review of Ecology and Systematics*, 30, 539-563
- Emshwiller E. & Doyle J.J. (2002) Origins of domestication and polyploidy in oca (*Oxalis Tuberosa* : Oxalidaceae). 2. Chloroplast-expressed glutamine synthetase data. *American Journal of Botany*, 89, 1042-1056
- Emshwiller E. (2006) Origins of polyploid crops - The example of the octoploid tuber crop *Oxalis tuberosa*. In: *Documenting domestication: New genetic and archaeological paradigms* (eds. Zeder MA, Bradley DG, Emshwiller E & Smith BD), pp. 153-168. University of California Press, Berkeley, CA
- Emshwiller E., Theim T., Grau A., Nina V. & Terrazas F. (2009) Origins of Domestication and Polyploidy in Oca (*Oxalis Tuberosa*; Oxalidaceae). 3. Aflp Data of Oca and Four Wild, Tuber-Bearing Taxa. *American Journal of Botany*, 96, 1839-1848
- Engle F. (1970) Exploration of the Chilca canyon, Peru. *Current Anthropology*, 11, 55 - 58
- Erickson D.L., Smith B.D., Clarke A.C., Sandweiss D.H. & Tuross N. (2005) An Asian origin for a 10,000-year-old domesticated plant in the Americas. *Proceedings of the National Academy of Science USA*, 102, 18315-20
- Esselink G.D., Nybom H. & Vosman B. (2004) Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting-peak ratios) method. *Theoretical and Applied Genetics*, 109, 402-408
- Evanno G., Regnaut S. & Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14, 2611-2620



- Excoffier L., Laval G. & Schneider S. (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1, 47-50
- Fajardo D.S., Bonte D.R.L. & Jarret R.L. (2002) Identifying and selecting for genetic diversity in Papua New Guinea sweetpotato *Ipomoea batatas* (L.) Lam. germplasm collected as botanical. *Genetic Resources and Crop Evolution*, 463-470
- Falush D., Stephens M. & Pritchard J.K. (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, 7, 574-578
- FAO (2011) Le Deuxième Rapport sur l'état des ressources phytogénétiques pour l'alimentation et l'agriculture dans le monde. In:
- Fernandez H., Hughes S., Vigne J.-D., Helmer D., Hodgins G., Miquel C., Hanni C., Luikart G. & Taberlet P. (2006) Divergent mtDNA lineages of goats in an Early Neolithic site, far from the initial domestication areas. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 15375-15379
- Fisher R. (1930) *The genetical theory of natural selection*. Clarendon Press, Oxford
- Fortune P.M., Schierenbeck K.A., Ainouche A.K., Jacquemin J., Wendel J.F. & Ainouche M.L. (2007) Evolutionary dynamics of Waxy and the origin of hexaploid *Spartina* species (Poaceae). *Molecular Phylogenetics and Evolution*, 43, 1040-1055
- Fregene M.A., Suarez M., Mkumbira J., Kulembeka H., Ndedya E., Kulaya A., Mitchel S., Gullberg U., Rosling H., Dixon A.G.O., Dean R. & Kresovich S. (2003) Simple sequence repeat marker diversity in cassava landraces: genetic diversity and differentiation in an asexually propagated crop. *Theoretical and Applied Genetics*, 107, 1083-1093
- Freitas F.O., Bendel G., Allaby R.G. & Brown T.A. (2003) DNA from primitive maize landraces and archaeological remains: implications for the domestication of maize and its expansion into South America. *Journal of Archaeological Science*, 30, 901-908
- Fuller D.Q. (2007) Contrasting patterns in crop domestication and domestication rates: Recent archaeobotanical insights from the old world. *Annals of Botany*, 100, 903-924
- Fuller D.Q. (2012) New archaeobotanical information on plant domestication from Macro-remains: tracking the evolution of domestication syndrome traits. In: *Biodiversity in agriculture: domestication, evolution, and sustainability* (eds. Gepts P, Famula TR, Bettinger RL, Brush SB, Damania AB, McGuire PE & Qualset CO), pp. 311-329. Cambridge University Press, Cambridge
- Gao M., Ashu G.M., Stewart L., Akwe W.a., Njiti V. & Barnes S. (2012) Wx intron variations support an allohexaploid origin of the sweetpotato [*Ipomoea batatas* (L.) Lam]. *Euphytica*, 177, 111-133
- Gebremedhin B., Ficetola G.F., Naderi S., Rezaei H.R., Maudet C., Rioux D., Luikart G., Flagstad O., Thuiller W. & Taberlet P. (2009) Frontiers in identifying conservation units: from neutral markers to adaptive genetic variation. *Animal Conservation*, 12, 107-109
- Gepts P., Osborn T.C., Rashka K. & Bliss F.A. (1986) Phaseolin-Protein Variability in Wild Forms and Landraces of the Common Bean (*Phaseolus-Vulgaris*) - Evidence for Multiple Centers of Domestication. *Economic Botany*, 40, 451-468
- Gepts P. (1998) Origin and evolution of common bean: past events and recent trends. *Horticultural Science*, 33, 1124-1130
- Gepts P. (2004) Crop domestication as a long-term selection experiment. *Plant Breeding Review*, 24, 1-44
- Gepts P F.T., Bettinger RL, Brush SB, Damania AB, McGuire PE, Qualset CO (2012) *Biodiversity in agriculture: domestication, evolution, and sustainability*. Cambridge University Press, Cambridge.
- Ghiselin M.T. (1975) A radical solution to the species problem. *Systematical Zoology*, 23, 536-544
- Gichuki S.T., Berenyi M., Zhang D., Hermann M., Glossl J. & Burg K. (2003) Genetic diversity in sweetpotato [*Ipomoea batatas* (L.) Lam.] in relationship to geographic sources as assessed with RAPD markers. *Genetic Resources and Crop Evolution*, 50, 429-437
- Glémin S. & Bataillon T. (2009) A comparative view of the evolution of grasses under domestication. *The New phytologist*, 183, 273-90

- Golson J. (1977) No room at the top: agricultural intensification in the New Guinea Highlands. In: *Sunda and Sahul* (eds. Allen J, Golson J & Jones R). Academic Press, London
- Golson J. & Hughes P.J. (1980) The Appearance of Plant and Animal Domestication in New Guinea. *Journal de la Société des océanistes*, 69, 294-303
- Golson J. (2007) Unravelling the story of early plant exploitation in highland Papua New Guinea. In: *Rethinking agriculture - Archaeological and ethnoarchaeological perspectives* (eds. Denham T, Iriarte J & Vrydaghs L), pp. 109 - 125. Left Coast Press, Walnut Creel, Cal.
- Gongora J., Rawlence N.J., Mobegi V.A., Jianlin H., Alcalde J.A., Matus J.T., Hanotte O., Moran C., Austin J.J., Ulm S., Anderson A.J., Larson G. & Cooper A. (2008) Indo-European and Asian origins for Chilean and Pacific chickens revealed by mtDNA (vol 105, pg 10308, 2008). *Proceedings of the National Academy of Sciences of the United States of America*, 105, 14234-14234
- Gorecki P.P. (1986) Human occupation and agricultural development in the Papua New Guinea Highlands. *Mountain Research and Development*, 6, 159-166
- Grant V. (1981) *Plant speciation, 2nd edition*. Columbia University Press, New York.
- Green R.C. (2003) Pacific Archaeology: Assessments and Anniversary of the First Lapita Excavation (July 1952). In: (ed. Sand C), pp. 95-120
- Green R.C. (2005) Sweet potato transfers in Polynesian prehistory. In: *The Sweet Potato in Oceania: a reappraisal*, pp. 43-62. Oceania Publications, Sydney
- Grivet L., Daniels C., Glaszmann J.C. & D'Hont A. (2004) A review of recent molecular genetics evidence for sugarcane evolution and domestication. *Ethnobotany research and applications*, 2, 9-17
- Gunn B.F., Baudouin L. & Olsen K.M. (2011) Independent Origins of Cultivated Coconut (*Cocos nucifera* L.) in the Old World Tropics. *Plos One*, 6
- Haberle S.G. & Atkin G. (2005) Needles in haystack: searching for sweet potato (*Ipomoea batatas* (L.) Lam.) in the fossil pollen record. In: *The Sweet Potato in Oceania: a reappraisal*, pp. 25 - 33. Oceania Publications, Sydney
- Handy E.S.C. (1940) *The Hawaiian Planter. Volume I: His plants, methods and areas of cultivation*. Bernice P. Bishop Museum, Honolulu.
- Harlan J.R. (1971) Agricultural origins: centers and noncenters. *Science*, 174, 468-74
- Harlan J.R. (1992) *Crops and man*. Am. Soc. Agronomy, Madison.
- Harris D.R. (2007) Agriculture, cultivation and domestication : xploring the conceptual framework of early food production. In: *Rethinking agriculture - Archaeological and ethnoarchaeological perspectives* (eds. Denham T, Iriarte J & Vrydaghs L), pp. 16-35. Left Coast Press, Walnut Creel, Cal.
- Hartigan J.A. & Wong M.A. (1979) A K-means clustering algorithm. *Applied Statistics*, 28, 100-108
- Harvey C., Gill G., Crossman C. & Fraser L.G. (1997) Assessing relationships of kumara cultivars by RAPD analysis. *New Zealand Journal of Botany*, 35, 479-485
- Hather J. & Kirch P.V. (1991) Prehistoric Sweet-Potato (*Ipomoea-Batatas*) from Mangaia Island, Central Polynesia. *Antiquity*, 65, 887-893
- Hays T.E. (1993) "The New Guinea Highlands": Region, culture area, or fuzzy set? . *Current Anthropology*, 34, 141-64
- Heerwaarden J.V., Doebley J., Briggs W.H., Glaubitz J.C., Goodman M.M., Sanchez J.d.J. & Ross-Ibarra J. (2011) Genetic signals of origin, spread and introgression in a large sample of maize landraces. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 1088-1092
- Helsen P., Verdyck P., Tye A. & Van Dongen S. (2009) Low levels of genetic differentiation between *Opuntia echios* varieties on Santa Cruz (Galapagos). *Plant Systematics and Evolution*, 279, 1-10
- Heun M., SchaferPregl R., Klawan D., Castagna R., Accerbi M., Borghi B. & F.Salamini (1997) Site of einkorn wheat domestication identified by DNA fingerprinting. *Science*, 278, 1312-1314.
- Hewlett B.S., Silvestri A.D. & Guglielmino C.R. (2002) Semes and genes in Africa. *Current Anthropology*, 43, 313-321

- Hijmans R.J.L., Guarino L., Bussink C., Mathur P., Cruz C., Barrientes I. & Rojas E. (2004) DIVA-GIS Vsn. 5.0. A geographic information system for the analysis of species distribution data. Manual available at <http://www.diva-gis.org>.
- Hillman G. & Davies M. (1990) Domestication rates in wild-type wheats and barley under primitive cultivation. *Biological Journal of the Linnean Society*, 39, 39-78
- Hilu K.W. (1993) Polyploidy and the evolution of domesticated plants. *American Journal of Botany*, 80, 1494-1499
- Hinkle A.E. (2007) Population structure of Pacific *Cordyline fruticosa* (Laxmanniaceae) with implications for human settlement of Polynesia. *American Journal of Botany*, 94, 828-839
- Hoffmann A.A. & Willi Y. (2008) Detecting genetic responses to environmental change. *Nature Review of Genetic*, 9
- Honne B. & Heun M. (2009) On the domestication genetics of self-fertilizing plants. *Vegetation History and Archaeobotany*, 18, 269-272
- Horrocks M., Shane P.a., Barber I.G., D'Costa D.M. & Nichol S.L. (2004) Microbotanical remains reveal Polynesian agriculture and mixed cropping in early New Zealand. *Review of Palaeobotanical Palynology*, 131, 147-157
- Horrocks M. & Rechtman R.B. (2009) Sweet potato (*Ipomoea batatas*) and banana (*Musa* sp.) microfossils in deposits from the Kona Field System, Island of Hawaii. *Journal of Archaeological Science*, 36, 1115-1126
- Huang J. & Sun M. (2000) Genetic diversity and relationships of sweetpotato and its wild relatives in *Ipomoea* series *Batatas* (Convolvulaceae) as revealed by inter-simple sequence repeat (ISSR) and restriction analysis of chloroplast DNA. *Theoretical and Applied Genetics*, 100, 1050-1060
- Huang J., Corke H. & Sun M. (2002) Highly polymorphic AFLP markers as a complementary tool to ITS sequences in assessing genetic diversity and phylogenetic relationships of sweetpotato (*Ipomoea batatas* (L.) Lam.) and its wild relatives. *Genetic Resources and Crop Evolution*, 49, 541-550
- Huelsenbeck J.P. & Ronquist F. (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17, 754-755
- Hughes C.E., Govindarajulu R., Robertson A., Filer D.L., Harris S.A. & Bailey C.D. (2007) Serendipitous backyard hybridization and the origin of crops. *Proceedings of the National Academy of Science of the United States of America*, 104, 14389-94
- Hurles M.E., Maund E., Nicholson J., Bosch E., Renfrew C., Sykes B.C. & Jobling M.A. (2003) Native American Y chromosomes in Polynesia: The genetic impact of the Polynesian slave trade. *American Journal of Human Genetics*, 72, 1282-1287
- Huson D.H. & Bryant D. (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23, 254-267
- Imazio S., Labra M., Grassi F., Scienza A. & Failla O. (2006) Chloroplast microsatellites to investigate the origin of grapevine. *Genetic Resources and Crop Evolution*, 53, 1003-1011
- Ingvarsson P.K., Ribstein S. & Taylor D.R. (2003) Molecular evolution of insertions and deletion in the chloroplast genome of *Silene*. *Molecular Biology and Evolution*, 20, 1737-1740
- Irwin G. (2006) Voyaging and settlement. In: *Vaka Moana: Voyages of the Ancestors* (ed. Howe KR). Bateman., Auckland
- Irwin G. (2011) Sailing from Polynesia to the Americas. In: *Polynesians in America: Pre-Columbian Contacts with the New World* (eds. Jones TL, Storey AA, Matisoo-Smith EA & Ramírez-Aliaga JM), pp. 247 - 262. Altamira Press, Plymouth
- Jaenicke-Despres V., Buckler E.S., Smith B.D., Gilbert M.T.P., Cooper A., Doebley J. & Paabo S. (2003) Early allelic selection in maize as revealed by ancient DNA. *Science*, 302, 1206-1208
- Jarret R.L., Gawel N. & Whitemore A. (1992) Phylogenetic-Relationships of the Sweet-Potato [*Ipomoea-Batatas* (L) Lam]. *Journal of the American Society for Horticultural Science*, 117, 633-637
- Jarret R.L. & Austin D.F. (1994) Genetic diversity and systematic relationships in sweetpotato (*Ipomoea batatas* (L.) Lam.) and related species as revealed by RAPD analysis. *Genetic Resources and Crop Evolution*, 41, 165-173

- Joly S. & Bruneau A. (2004) Evolution of triploidy in *Apios americana* (Leguminosae) revealed by genealogical analysis of the histone H3-D gene. *Evolution*, 58, 284-295
- Jombart T. (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403-1405
- Jombart T., Devillard S. & Balloux F. (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, 11
- Jones A. (1965) Cytological observations and fertility measurements of sweetpotato (*Ipomoea batatas* (L.) Lam.). *Proceedings of the American Society of Horticulture Science*, 86, 527-537
- Jones A. & Deonier M.T. (1965) Interspecific crosses among *Ipomoea lacunosa*, *I. ramosa*, *I. trichocarpa* and *I. triloba*. *Botanical Gazette*, 126, 226-232
- Jones A. (1967) Should Nishiyama 's K123 (*Ipomoea trifida*) Be Designated *I. batatas* ? *Economic Botany*, 21, 163-166
- Jones T.L. (2011) Re-introducing the case for polynesian contact. In: *Polynesians in America: Pre-Columbian Contacts with the New World* (eds. Jones TL, Storey AA, Matisoo-Smith EA & Ramirez-Aliaga JM), pp. 1 - 6. Altamira Press, Plymouth
- Jones T.L., Storey A.A., Matisoo-Smith E.A. & Ramirez-Aliaga J.M. (2011) *Polynesians in America: Pre-Columbian Contacts with the New World* Altamira Press, Plymouth.
- Jordan F.M., Gray R.D., Greenhill S.J. & Mace R. (2009) Matrilocal residence is ancestral in Austronesian societies. *Proceedings B of the Royal Society*, 276, 1957-1964
- Kaplan L. & Lynch T.F. (1999) *Phaseolus* (Fabaceae) in archaeology: AMS radiocarbon dates and their significance for pre-Columbian agriculture. *Economic Botany*, 53, 261-272
- Kayser M. (2010) The human genetic history of Oceania: near and remote views of dispersal. *Current Biology*, 20, R194-201
- Kilian B., Ozkan H., Walther A., Kohl J., Dagan T., Salamini F. & Martin W. (2007) Molecular diversity at 18 loci in 321 wild and 92 domesticate lines reveal no reduction of nucleotide diversity during *Triticum monococcum* (einkorn) domestication: Implications for the origin of agriculture. *Molecular Biology and Evolution*, 24, 2657-2668
- Kirch P. (2002) *On the Road of the Winds - An Archaeological History of the Pacific Islands before European Contact*. University of California Press, Berkeley.
- Kirch P.V. & Rallu J.L. (2007) *The growth and collapse of Pacific island societies: archaeological and demographic perspectives*. University of Hawaii Press, Honolulu.
- Kirch P.V. (2010) Peopling of the Pacific: A Holistic Anthropological Perspective. In: *Annual Review of Anthropology, Volume 39*, pp. 131-148
- Knapp M., Clarke A.C., Horsburgh K.A. & Matisoo-Smith E.A. (2012) Setting the stage - Building and working in an ancient DNA laboratory. *Annals of Anatomy-Anatomischer Anzeiger*, 194, 3-6
- Kobayashi M. (1984) The *Ipomoea trifida* complex closely related to sweet potato. In: *Proceedings of the 6th Symposium of the International Society of Tropical Root Crop* (eds. Shideler SF & Rincon H), pp. 561-568. CIP, Lima, Peru
- Komaki K., Regmi H., Katayama K. & Tamiya S. (1998) Morphological and RAPD pattern variations in sweetpotato and its closely related species. *Breeding Science*, 48, 281-286
- Komatsuda T., Pourkheirandish M., He C., Azhaguvel P., Kanamori H., Perovic D., Stein N., Graner A., Wicker T., Tagiri A., Lundqvist U., Fujimura T., Matsuoka M., Matsumoto T. & Yano M. (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 1424-1429
- Kosman E. & Leonard K.J. (2005) Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. *Molecular Ecology*, 14, 415-424
- Kowiyama Y., Shimano N. & Kawase T. (1980) Genetic-Analysis of Incompatibility in the Diploid *Ipomoea* Species Closely Related to the Sweet-Potato. *Theoretical and Applied Genetics*, 58, 149-155

- Kreike C.M., Van Eck H.J. & Lebot V. (2004) Genetic diversity of taro, *Colocasia esculenta* (L.) Schott, in Southeast Asia and the Pacific ok. *Theoretical and Applied Genetics*, 109, 761-768
- Kuhner M.K. (2009) Coalescent genealogy samplers: windows into population history. *Trends in Ecology & Evolution*, 24, 86-93
- Ladefoged T.N., Graves M.W. & James H.C. (2005) The introduction of sweet potato in Polynesia : early remains in Hawai'i. *Journal of Polynesian Society*, 114, 359-374
- Larson G., Albarella U., Dobney K., Rowley-Conwy P., Schibler J., Tresset A., Vigne J.D., Edwards C.J., Schlumbaum A., Dinu A., Balacescu A., Dolman G., Tagliacozzo A., Manaseryan N., Miracle P., Van Wijngaarden-Bakker L., Masseti M., Bradley D.G. & Cooper A. (2007a) Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proceedings of the National Academy of Sciences*, 104, 15276-81
- Larson G., Cucchi T., Fujita M., Matisoo-Smith E., Robins J., Anderson A., Rolett B., Spriggs M., Dolman G., Kim T.-H., Thuy N.T.D., Randi E., Doherty M., Due R.A., Bollt R., Djubiantono T., Griffin B., Intoh M., Keane E., Kirch P., Li K.-T., Morwood M., PedriÃ±a L.M., Piper P.J., Rabett R.J., Shooter P., Van den Bergh G., West E., Wickler S., Yuan J., Cooper A. & Dobney K. (2007b) Phylogeny and ancient DNA of *Sus* provides insights into neolithic expansion in Island Southeast Asia and Oceania. *Proceedings of the National Academy of Sciences*, 104, 4834-9
- Larson G., Liu R., Zhao X., Yuan J., Fuller D., Barton L., Dobney K., Fan Q., Gu Z., Liu X.-H., Luo Y., Lv P., Andersson L. & Li N. (2010) Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. *Proceedings of the National Academy of Sciences*, 107, 7686-91
- Lawler A. (2010) Beyond Kon-Tiki: did Polynesians sail to South America? . *Science* 328, 1344-1347
- Leach H. (2005) *Ufi kmura*, the sweet potato as yam. In: *The Sweet Potato in Oceania: a reappraisal*, pp. 63 - 70. Oceania Publications, Sydney
- Lebot V. & Aradhya K.M. (1991) Isozyme Variation in Taro (*Colocasia-Esculenta* (L.) Schott) from Asia and Oceania. *Euphytica*, 56, 55-66
- Lebot V. (1992) Genetic Vulnerability of Oceanias Traditional Crops. *Experimental Agriculture*, 28, 309-323
- Lebot V., Merlin M. & Lindstrom L. (1997) *Kava, the Pacific elixir: The definitive guide to its ethnobotany, history and chemistry*. Healing arts press, Rochester.
- Lebot V. (1999a) Biomolecular evidence for plant domestication in Sahul. *Genetic Resources and Crop Evolution*, 46, 619-628
- Lebot V. (1999b) Biomolecular evidence for plant domestication in Sahul. *Genetic Resources and Crop Evolution*, 46, 619-628
- Lebot V., Prana M.S., Kreike N., van Heck H., Pardales J., Okpul T., Gendua T., Thongjiem M., Hue H., Viet N. & Yap T.C. (2004) Characterisation of taro (*Colocasia esculenta* (L.) Schott) genetic resources in Southeast Asia and Oceania. *Genetic Resources and Crop Evolution*, 51, 381-392
- Lebot V., Ivancic A. & Abraham K. (2005) The geographical distribution of allelic diversity, a practical means of preserving and using minor root crop genetic resources. *Experimental Agriculture*, 41, 475-489
- Lebot V. (2009a) *Tropical Root and Tuber Crops - Cassava, Sweet Potato, Yams and Aroids*. CABI Publishing Group, Wallingford, UK.
- Lebot V. (2009b) Tropical Root and Tuber Crops: Cassava, Sweet Potato, Yams, Aroids. In, p. 432
- Lee T., Burch J.B., Coote T., Fontaine B., Gargominy O., Pearce-Kelly P. & Foighil D.O. (2007) Prehistoric inter-archipelago trading of Polynesian tree snails leaves a conservation legacy. *Proceedings of the Royal Society B-Biological Sciences*, 274, 2907-2914
- Lenné J.M., Sweetmore A. & Burstow A. (1994) Morphological and pathogenic characterisation of *Elsinoe batatas*: causal agent of sweetpotato scab. In: *Proceedings of the International Congress for Plant Protection in the Tropics*, pp. 64-66, Kuala Lumpur, Malaysia
- Leotard G., Duputie A., Kjellberg F., Douzery E.J., Debain C., de Granville J.J. & McKey D. (2009) Phylogeography and the origin of cassava: new insights from the northern rim of the Amazonian basin. *Molecular Phylogenetic and Evolution*, 53, 329-34

- Levin D.A. (2002) *The role of chromosomal change in plant evolution*. Oxford University Press, Oxford.
- Lewthwhite S.L. (2004) Storage root production in sweet potatoe (*Ipomoea batatas* (L.) Lam.). In: Massey University, Palmerston North, New-Zealand
- Lie B.A., Dupuy B.M., Spurkland A., Fernandez-Vina M.A., Hagelberg E. & Thorsby E. (2007) Molecular genetic studies of natives on Easter Island: evidence of an early European and Amerindian contribution to the Polynesian gene pool. *Tissue Antigens*, 69, 10-18
- Liu A. & Burke J.M. (2006) Patterns of nucleotide diversity in wild and cultivated sunflower. *Genetics*, 173, 321-30
- Londo J.P., Chiang Y.C., Hung K.H., Chiang T.Y. & Schaal B.A. (2006) Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 9578-9583
- Lum J.K., Kaneko A., Tanabe K., Takahashi N., Bjorkman A. & Kobayakawa T. (2004) Malaria dispersal among islands: human mediated *Plasmodium falciparum* gene flow in Vanuatu, Melanesia. *Acta Tropica*, 90, 181-185
- Lutulele R.P. (2001) Sweet potato variety developments in the PNG highlands: implications for future research and extension focus. In: *Food Security for Papua New Guinea, Proceedings of the Papua New Guinea Food and Nutrition 2000 Conference (ACIAR Proceedings 99)* (eds. Bourke RM, Allen MG & Salisbury JD), pp. 683-688, University of Technology: Lae
- Magoon M.L., Krishnan R. & Vijaya Bai K. (1970) Cytological evidence on the origin of sweet potato. *Theoretical and Applied Genetics*, 40, 360-366
- Malapa R., Arnau G., Noyer J.L. & Lebot V. (2005) Genetic diversity of the greater yam (*Dioscorea alata* L.) and relatedness to *Dioscorea nummularia* Lam. and *Dioscorea transversa* Br. as revealed with AFLP markers. *Genetic Resources and Crop Evolution*, 52, 919-929
- Mallet J. (1995) A Species Definition for the Modern Synthesis. *Trends in Ecology & Evolution*, 10, 294-299
- Martin F.W. & Jones A. (1973) The species of *Ipomoea* closely related to the sweetpotato. *Economic Botany*, 201-215
- Matisoo-Smith E., Roberts R.M., Irwin G.J., Allen J.S., Penny D. & Lambert D.M. (1998) Patterns of prehistoric human mobility in Polynesia indicated by mtDNA from the Pacific rat. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 15145-15150
- Matisoo-Smith E. (2002) Something old, something new: Do genetic studies of contemporary populations reliably represent prehistoric populations of Pacific *Rattus exulans*? *Human Biology*, 74, 489-496
- Matisoo-Smith E., Hingston M., Summerhayes G., Robins J., Ross H.A. & Hendy M. (2009) On the Rat Trail in Near Oceania: Applying the Commensal Model to the Question of the Lapita Colonization. *Pacific Science*, 63, 465-475
- Matisoo-Smith E. & Robins J. (2009) Mitochondrial DNA evidence for the spread of Pacific rats through Oceania. *Biological Invasions*, 11, 1521-1527
- Matsuoka Y., Vigouroux Y., Goodman M.M., {Sanchez G} J., Buckler E. & Doebley J. (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences USA*, 99, 6080-4
- Mayr E. (1942) *Systematics and the origin of species from the viewpoint of a zoologist*. Columbia University Press, New York.
- Mayr E. (2000) The biological species concept. In: *Species concepts and phylogenetic theory* (eds. Wheeler QD & Meier R). Columbia University Press, New York
- McCouch S., Kovach M., Sweeney M., Jiang H. & Semon M. (2012) The dynamics of rice domestication: a balance between gene flow and genetic isolation. In: *Biodiversity in agriculture: domestication, evolution, and sustainability* (eds. Gepts P, Famula TR, Bettinger RL, Brush SB, Damania AB, McGuire PE & Qualset CO), pp. 311-329. Cambridge University Press, Cambridge
- McDonald J.A. & Austin D.F. (1990) Changes and additions in *Ipomoea* section *Batatas* (Convolvulaceae). *Brittonia*, 42, 116-120

- McKay J.K. & Latta R.G. (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution*, 17, 285-291
- McKey D., Elias M., Pujol B. & Duputié A. (2010) The evolutionary ecology of clonally propagated domesticated plants. *New Phytologist*, 186, 318-332
- McKey D., Elias M., Pujol B. & Duputié A. (2012) Ecological approaches to crop domestication. In: *Biodiversity in agriculture: domestication, evolution, and sustainability* (eds. Gepts P, Famula TR, Bettinger RL, Brush SB, Damania AB, McGuire PE & Qualset CO), pp. 377-406. Cambridge University Press, Cambridge
- Meilleur B.A. (1998) Clones within clones: cosmology and esthetics in Polynesian crop selection. *Anthropologica*, 40, 71-82
- Miller A.J. & Gross B.L. (2011) From forest to field: perennial fruit crop domestication. *American Journal of Botany*, 98, 1389-414
- Mishler B.D. & Theriot E.C. (2000) The phylogenetic species concept (sensu Mishler and Theriot): monophyly, apomorphy, and phylogenetic species concepts. In: *Species concepts and phylogenetic theory* (eds. Wheeler QD & Meier R), pp. 44-54. Columbia University Press, New York
- Moodley Y., Linz B., Yamaoka Y., Windsor H.M., Breurec S., Wu J.Y., Maady A., Bernhoft S., Thiberge J.M., Phuanukoonnon S., Jobb G., Siba P., Graham D.Y., Marshall B.J. & Achtman M. (2009) The Peopling of the Pacific from a Bacterial Perspective. *Science*, 323, 527-530
- Morrell P.L. & Clegg M.T. (2007) Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 3289-3294
- Motta-Aldana J.R., Serrano-Serrano M.L., Hernández-Torres J., Castillo-Villamizar G., Debouck D.G. & Chacón S. M.I. (2010) Multiple origins of Lima bean landraces in the Americas: Evidence from chloroplast and nuclear DNA polymorphisms. *Crop Science*, 50, 1773-1787
- Nakanishi T. & Kobayashi N. (1979) Geographic distribution of cross incompatibility group in sweet potato. *Incompatibility Newsletter*, 11, 72-75
- Nichols R. (2001) Gene trees and species trees are not the same. *Trends in Ecology & Evolution*, 16, 358-364
- Nishiyama I., Miyazaki T. & Sakamoto S. (1975) Evolutionary Autopoloidy in Sweet-Potato (*Ipomoea-Batatas* (L) Lam) and Its Progenitors. *Euphytica*, 24, 197-208
- Nunn P.D. (2000) Illuminating sea-level fall around AD 1220-1510 (730-440 cal yr BP) in the Pacific Islands: implications for environmental change and cultural transformation. *New Zealand Geographer*, 56, 46-54
- Nunn P.D. (2003) Revising ideas about environmental determinism: Human-environment relations in the Pacific Islands. *Asia Pacific Viewpoint* 44, 63-72
- Nunn P.D. (2009) Responding to the challenges of climate change in the Pacific Islands: management and technological imperatives. *Climate Research*, 40, 211-231
- Obbard D.J., Harris S.A. & Pannell J.R. (2006) Simple allelic-phenotype diversity and differentiation statistics for allopolyploids. *Heredity*, 97, 296-303
- Ollitrault P. & Luro F. (2001) Citrus. In: *Tropical Plant Breeding* (ed. Charrier A JM, Hamon S, Nicolas D), pp. 55-77. CIRAD, Paris
- Olsen K.M. & Schaal B.A. (1999) Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 5586-5591
- Olsen K.M. & Schaal B.A. (2001) Microsatellite variation in cassava (*Manihot esculenta*, Euphorbiaceae) and its wild relatives: Further evidence for a southern Amazonian origin of domestication. *American Journal of Botany*, 88, 131-142
- Olsen K.M. (2004) SNPs, SSRs and inferences on cassava's origin. *Plant Molecular Biology*, 56, 517-526
- Oracion M.Z., Niwa K. & Shiotani I. (1990) Cytological Analysis of Tetraploid Hybrids between Sweet-Potato and Diploid *Ipomoea trifida* (Hbk) Don. *Theoretical and Applied Genetics*, 80, 617-624

- Orjeda G., Freyre R. & Iwanaga M. (1991) Use of *Ipomoea trifida* germ plasm for sweet potato improvement. 3. Development of 4x interspecific hybrids between *Ipomoea batatas* (L.) Lam ( $2n=6x=90$ ) and *I. trifida* (H.B.K) G. Don ( $2n=2x=30$ ) as storage-root initiators for wild species. *Theoretical and Applied Genetics*, 83, 159-163
- Otto S.P. (2007) The evolutionary consequences of polyploidy. *Cell*, 131, 452-462
- Oumar I., Mariac C., Pham J.-L. & Vigouroux Y. (2008) Phylogeny and origin of pearl millet (*Pennisetum glaucum* [L.] R. Br) as revealed by microsatellite loci. *Theoretical and Applied Genetics*, 117, 489-497
- Paradis E., Claude J. & Strimmer K. (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20, 289-290
- Parisod C. & Besnard G. (2007) Glacial in situ survival in the Western Alps and polytopic autopolyploidy in *Biscutella laevigata* L.(Brassicaceae). *Molecular Ecology*, 16, 2755-2767
- Parisod C., Holderegger R. & Brochmann C. (2010) Evolutionary consequences of autopolyploidy. *New Phytologist*, 186, 5-17
- Paterson A.H., Lin Y.R., Li Z.K., Schertz K.F., Doebley J.F., Pinson S.R.M., Liu S.C., Stansel J.W. & Irvine J.E. (1995) Convergent Domestication of Cereal Crops by Independent Mutations at Corresponding Genetic-Loci. *Science*, 269, 1714-1718
- Paterson H.E.H. (1985) The recognition concept of species. In: *Species and speciation* (ed. Vrba ES), pp. 21-29. Transvaal Museum, Pretoria
- Paun O., Schonswetter P., Winkler M. & Tribsch A. (2008) Historical divergence vs. contemporary gene flow: evolutionary history of the calcicole *Ranunculus alpestris* group (Ranunculaceae) in the European Alps and the Carpathians. *Mol Ecol*, 17, 4263-75
- Peakall R. & Smouse P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288-295
- Perales H.R., Benz B.F. & Brush S.B. (2005) Maize diversity and ethnolinguistic diversity in Chiapas, Mexico. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 949-954
- Perrier X., Flori A. & Bonnot F. (1999) Les méthodes d'analyse des données. In: *Diversité génétique des plantes tropicales cultivées* (eds. Hamon P, Seguin M, Perrier X & Glaszmann JC), pp. 43-87. CIRAD, Montpellier
- Perrier X. & Jacquemoud-Collet J.P. (2006) DARwin software In, p. <http://darwin.cirad.fr/>
- Perrier X., De Langhe E., Donohue M., Lentfer C., Vrydaghs L., Bakry F., Carreel F., Hippolyte I., Horry J.-P., Jenny C., Lebot V., Risterucci A.-M., Tomekpe K., Doutrelepont H., Ball T., Manwaring J., de Maret P. & Denham T. (2011) Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proceedings of the National Academy of Sciences USA*, 108, 11311-11318
- Petit R.J., El Mousadik A. & Pons O. (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, 12, 844 – 855
- Petit R.J., Duminil J., Fineschi S., Hampe A., Salvini D. & Vendramin G.G. (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, 14, 689-701
- Petit R.J. & Excoffier L. (2009) Gene flow and species delimitation. *Trends in Ecology & Evolution*, 24, 386-393
- Pickersgill B. (1989) Cytological and genetical evidence on the domestication and diffusion of crops within the Americas. In: *Foraging and farming: the evolution of plant exploitation* (eds. Harris DR & Hillman GC), pp. 426-439. Unwin Hyman, London
- Pickersgill B. (2004) Crops and cultures in the pacific : new data and new techniques for the investigation of old questions. *Ethnobotany research and applications*, 2, 1 - 8
- Pickersgill B. (2007) Domestication of plants in the Americas: insights from Mendelian and molecular genetics. *Annals of Botany*, 100, 925-40
- Piperno D.R., Andres T.C. & Stothert K.E. (2000) Phytoliths in Cucurbita and other neotropical curcubitaceae and their occurrence in early archaeological sites from the Lowland American tropics. *Journal of Archaeological Science*, 27, 193-208
- Piperno D.R. (2011) The origins of plant cultivation and domestication in the New World tropics. *Current Anthropology*, 52, 453-470



- Posada D. & Crandall K.A. (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution*, 16, 37-45
- Price T.D. & Bar-Yosef O. (2011) The Origins of Agriculture: New Data, New Ideas An Introduction to Supplement 4. *Current Anthropology*, 52, S163-S174
- Pritchard J.K., Stephens M. & Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155, 945-959
- Provan J., Powell W., Dewar H., Bryan G., Machray G.C. & Waugh R. (1999) An extreme cytoplasmic bottleneck in the modern European cultivated potato (*Solanum tuberosum*) is not reflected in decreased levels of nuclear diversity. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 266, 633-639
- Provan J., Powell W. & Hollingsworth P.M. (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology & Evolution*, 16, 142-147
- Pujol B., Gigot G., Laurent G., Pinheiro-Kluppel M., Elias M., Hossaert-McKey M. & McKey D. (2002) Germination ecology of cassava (*Manihot esculenta* Crantz, Euphorbiaceae) in traditional agroecosystems: Seed and seedling biology of a vegetatively propagated domesticated plant. *Economic Botany*, 56, 366-379
- Pujol B., Muhlen G., Garwood N., Horoszowski Y., Douzery E.J. & McKey D. (2005) Evolution under domestication: contrasting functional morphology of seedlings in domesticated cassava and its closest wild relatives. *New Phytol*, 166, 305-18
- Pujol B. & McKey D. (2006) Size asymmetry in intraspecific competition and the density-dependence of inbreeding depression in a natural plant population: a case study in cassava (*Manihot esculenta* Crantz, Euphorbiaceae). *Journal of Evolutionary Biology*, 19, 85-96
- Pujol B., Renoux F., Elias M., Rival L. & McKey D. (2007) The unappreciated ecology of landrace populations: Conservation consequences of soil seed banks in Cassava. *Biological Conservation*, 136, 541-551
- Purugganan M.D. & Fuller D.Q. (2009) The nature of selection during plant domestication. *Nature*, 457, 843-848
- Rajapakse S., Nilmalgoda S.D., Molnar M., Ballard R.E., Austin D.F. & Bohac J.R. (2004) Phylogenetic relationships of the sweetpotato in *Ipomoea* series *Batatas* (Convolvulaceae) based on nuclear beta-amylase gene sequences. *Molecular Phylogenetics and Evolution*, 30, 623-632
- Ramírez-Aliaga J.M. (2011) The Mapuche connection. In: *Polynesians in America: Pre-Columbian Contacts with the New World* (eds. Jones TL, Storey AA, Matisoo-Smith EA & Ramírez-Aliaga JM), pp. 95 - 110. Altamira Press, Plymouth
- Ramsey J. & Schemske D.W. (1998) Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics*, 29, 467-501
- Ramsey J. & Schemske D.W. (2002) Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics*, 33, 589-639
- Ranere A.J., Piperno D.R., Holst I., Dickau R. & Iriarte J. (2009) The cultural and chronological context of early Holocene maize and squash domestication in the Central Balsas River Valley, Mexico. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 5014-8
- Ravi V., Naskar S., Makesh Kumar T., Babu B. & Krishnan B.S.P. (2009) Molecular physiology of storage root formation and development in sweet potato (*Ipomoea batatas* (L.) Lam.). *Journal of Root Crops*, 35, 1-27
- Rensch K.H. (1991) Linguistic Analysis and Ethnobotany, Expectations and Limitations. In: *Islands, Plants, and Polynesians* (eds. Cox PA & Banack S), pp. 97-111. Dioscorides Press, Portland, Oregon
- Richerson P.J., Boyd R. & Bettinger R.L. (2001) Was Agriculture Impossible during the Pleistocene but Mandatory during the Holocene? A Climate Change Hypothesis. *American Antiquity*, 66, 387-411
- Rieseberg L.H. & Willis J.H. (2007) Plant speciation. *Science*, 317, 910-914
- Rival L. & McKey D. (2008) Domestication and Diversity in Manioc (*Manihot esculenta* Crantz ssp *esculenta*, Euphorbiaceae). *Current Anthropology*, 49, 1116-1125

- Robertson A., Rich T.C.G., Allen A.M., Houston L., Roberts C., Bridle J.R., Harris S.A. & Hiscock S.J. (2010) Hybridization and polyploidy as drivers of continuing evolution and speciation in *Sorbus*. *Molecular ecology*, 19, 1675-1690
- Romao R.L. (2000) Northeast Brazil: A secondary center of diversity for watermelon (*Citrullus lanatus*). *Genetic Resources and Crop Evolution*, 47, 207-213
- Roullier C., Rossel G., Tay D., McKey D. & Lebot V. (2011) Combining chloroplast and nuclear microsatellites to investigate origin and dispersal of New World sweet potato landraces. *Molecular Ecology*, 20, 3963-3977
- Rozen S. & Skaletsky H.J. (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds. Krawetz S & Misener S), pp. 365-386. Humana Press, Totowa
- Salick J., Cellinese N. & Knapp S. (1997) Indigenous diversity of cassava: generation, maintenance, use and loss among the Amuesha, Peruvian upper Amazon. *Economic Botany*, 51, 6-19
- Sanjur O.I., Piperno D.R., Andres T.C. & Wessel-Beaver L. (2002) Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from a mitochondrial gene: Implications for crop plant evolution and areas of origin. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 535-40
- Sardos J., McKey D., Duval M.F., Malapa R., Noyer J.L. & Lebot V. (2008) Evolution of cassava (*Manihot esculenta* Crantz) after recent introduction into a South Pacific Island system: the contribution of sex to the diversification of a clonally propagated crop. *Genome*, 51, 912-921
- Sauer C.O. (1952) *Agricultural origins and dispersals*. American Geographical Society, New York.
- Scaglione R. & Soto K. (1994) A prehistoric introduction of sweet potato in New Guinea? . In: *Migration and Transformation: Regional Perspectives on New Guinea* (eds. Strathern AJ & Sturzenhofecker G), pp. 257-294. ASAO Monograph 15, University of Pittsburgh Press, Pittsburgh
- Scaglione R. (2005) *Kumara* in the Ecuadorian gulf of Guyadaquil? . In: *The Sweet Potato in Oceania: a reappraisal*, pp. 35-42. Ocenania publications, Sydney
- Scarcelli N. (2005) Structure et dynamique de la diversité d'une plante cultivée à multiplication végétative: le cas des ignames au Bénin (*Dioscorea* sp.). In, p. 221. Université Montpellier 2, Montpellier
- Scarcelli N., Tostain S., Vigouroux Y., Agbangla C., Dainou O. & Pham J.L. (2006) Farmers' use of wild relative and sexual reproduction in a vegetatively propagated crop. The case of yam in Benin. *Mol Ecol*, 15, 2421-31
- Schneider J. (1995) Farmer practices and sweet potato diversity in highland New Guinea. In: *Indigenous Knowledge in Conservation of Crop Genetic Resources : Proceedings of an International Workshop* (ed. Schneider J), pp. 63-70, Cisarua, Bogor, Indonesia
- Seelenfreund D., Clarke A.C., Oyanedel N., Pina R., Lobos S., Matisoo-Smith E.A. & Seelenfreund A. (2012) Paper mulberry (*Broussonetia papyrifera*) as a commensal model for human mobility in Oceania: anthropological, botanical and genetic considerations. *New Zealand Journal of Botany*, 48, 231-247
- Segraves K.A., Thompson J.N., Soltis P.S. & Soltis D.E. (1999) Multiple origins of polyploidy and the geographic structure of *Heuchera grossulariifolia*. *Molecular Ecology*, 8, 253-262
- Shaw J., Lickey E.B., Schilling E.E. & Small R.L. (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany*, 94, 275-288
- Shiotani I., Yoshida S. & Kawase T. (1990) Numerical taxonomic analysis and crossability of diploid *Ipomoea* species related to the sweet potato. . *Japanese Journal of Breeding*, 40, 159-174
- Simeoni P. & Lebot V. (2012) Spatial Representation of Land Use and Population Density: Integrated Layers of Data Contribute to Environmental Planning in Vanuatu. *Human Ecology*, 40, 541-555
- Simmons M.P. & Ochoterena H. (2000) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology*, 49, 369-381
- Singh S.P., Gepts P. & Debouck D.G. (1991) Races of Common Bean (*Phaseolus-Vulgaris*, Fabaceae). *Economic Botany*, 45, 379-396

- Smith B.D. (2006) Eastern North America as an independent center of plant domestication. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 12223-8
- Smith P.E.L. (1976) *Food production and its consequences*, Menlopark, Cal.
- Soltis D.E. & Soltis P.S. (1993) Molecular-Data and the Dynamic Nature of Polyploidy. *Critical Reviews in Plant Sciences*, 12, 243-273
- Soltis D.E. & Soltis P.S. (1999) Polyploidy: recurrent formation and genome evolution. *Trends in Ecology & Evolution*, 14, 348-352
- Soltis D.E., Soltis P.S. & Tate J. (2004) Advances in the study of polyploidy since *Plant speciation*. *New Phytologist*, 161, 173-191
- Spillane C. & Gepts P. (2001) Evolutionary and genetic perspectives on the dynamics of crop gene pools. In: *Broadening the genetic base of crop production* (eds. Cooper HD, Spillane C & Hodgkin T), pp. 25-70. CABI, Wallingford
- Spooner D.M., McLean K., Ramsay G., Waugh R. & Bryan G.J. (2005) A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 14694-14699
- Srisuwan S., Sihachakr D. & Siljak-Yakovlev S. (2006) The origin and evolution of sweet potato (*Ipomoea batatas* Lam.) and its wild relatives through the cytogenetic approaches. *Plant Science*, 171, 424-433
- Stebbins G.L. (1950) *Variation and evolution in plants*. Columbia University Press, New York.
- Stebbins G.L. (1971) *Chromosomal evolution in higher plants*. Edward Arnold, London.
- Stebbins G.L. (1980) Polyploidy in plants: unsolved problems and prospects. In: *Polyploidy: biological relevance* (ed. Lewis WH), pp. 495-520. Plenum Press
- Stokes J. (2007) Kumara claim becomes hot potato. *New Zealand Herald*
- Storey A.A., Ram  rez J.M., Quiroz D., Burley D.V., Addison D.J., Walter R., Anderson A.J., Hunt T.L., Athens J.S., Huynen L. & Matisoo-Smith E.A. (2007a) Radiocarbon and DNA evidence for a pre-Columbian introduction of Polynesian chickens to Chile. *Proceedings of the National Academy of Sciences*, 104, 10335-9
- Storey A.A., Ramirez J.M., Quiroz D., Burley D.V., Addison D.J., Walter R., Anderson A.J., Hunt T.L., Athens J.S., Huynen L. & Matisoo-Smith E.A. (2007b) Radiocarbon and DNA evidence for a pre-Columbian introduction of Polynesian chickens to Chile. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 10335-10339
- Storey A.A., Quiroz D., Ramirez J.M., Beavan-Athfield N., Addison D.J., Walter R., Hunt T., Athens J.S., Huynen L. & Matisoo-Smith E.A. (2008) Pre-Columbian chickens, dates, isotopes, and mtDNA. *Proceedings of the National Academy of Sciences of the United States of America*, 105, E100-E99; author reply E100
- Storey A.A. (2011) Identifying contact with the Americas: A commensal-based approach. In: *Polynesians in America: Pre-Columbian Contacts with the New World* (eds. Jones TL, Storey AA, Matisoo-Smith EA & Ram  rez-Aliaga JM), pp. 111 - 138. Altamira Press, Plymouth
- Tanno K. & Willcox G. (2006) How fast was wild wheat domesticated? *Science*, 311, 1886-1886
- This P., Lacombe T. & Thomas R. (2006) Historical origins and genetic diversity of wine grapes. *Trends in Genetics*, 22, 511-519
- Thorsby E. (2012) The Polynesian gene pool: an early contribution by Amerindians to Easter Island. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 367, 812-819
- Ting Y.C., Kehr A.E. & Miller J.C. (1957) A cytological study of the sweetpotato plant *Ipomoea batatas* (L.) Lam and its related species. *American Naturalist*, 91, 197-203
- Ugent D. & Peterson W. (1988) Archeological remains of potato and sweet potato in Peru. *CIP Circular*, 16, 1-10
- Van Dijk P. & Bakx-Schotman T. (1997) Chloroplast DNA phylogeography and cytotype geography in autopolyploid *Plantago media*. *Molecular Ecology*, 9, 345-352
- Van Heerwaarden J., Doebley J., Briggs W.H., Glaubitz J.C., Goodman M.M., Gonzalez J.D.S. & Ross-Ibarra J. (2010) Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 1088-1092

- Van Oostroom S.J. (1953) Convolvulaceae. In: *Flora malesiana, ser. 1, 4:(4)* (ed. van Steenis CFGJ), pp. 388-512
- Van Wijmeersch P. (2001) The status of sweet potato variety evaluation in PNG and recommendations for further research. In: *The Papua New Guinea Food and Nutrition 2000 Conference (ACIAR Proceedings 99)* (eds. Bourke RM, Allen MG & Salisbury JD), pp. 683-688, University of Technology: Lae
- Verdcourt B. (1963) Convolvulaceae. In: *Flora of tropical East Africa* (eds. Hubbard CE & Milne-Redhead E), p. 162, London
- Vigouroux Y., Glaubitz J.C., Matsuoka Y., Goodman M.M., Jesus S.G. & Doebley J. (2008) Population Structure and Genetic Diversity of New World Maize Races Assessed by DNA Microsatellites. *American Journal of Botany*, 95, 1240-1253
- Vrydaghs L. & Denham T. (2007) Rethinking agriculture : introductory thoughts. In: *Rethinking agriculture - Archaeological and ethnoarchaeological perspectives* (eds. Denham T, Iriarte J & Vrydaghs L), pp. 1 - 15. Left Coast Press, Walnut Creel, Cal.
- Wallin P., Stevenson C. & Ladefoged T. (2005) Sweet potato production on Rapa Nui. In: *The Sweet Potato in Oceania: a reappraisal*, pp. 85-88. Oceania Publications, Sydney
- Walter A. & Lebot V. (2007) *Gardens of Oceania*. IRD Editions - Cirad, Canberra.
- Wang R.-L., A. Stec, J. Hey, L. Lukens, and J. Doebley (1999) The limits of selection during maize domestication. *Nature*, 398, 236-239
- Wang R.-L., A. Stec, J. Hey, L. Lukens, and J. Doebley (2001) Correction: The limits of selection during maize domestication. *Nature*, 410, 718
- Weissmann S., Feldman M., Gressel J. & Gressel J. (2005) Sequence evidence for sporadic intergeneric DNA introgression from wheat into a wild Aegilops species. *Molecular Biology and Evolution*, 22, 2055-2062
- Wendel J.F., Brubaker C.L. & Percival A.E. (1992) Genetic diversity in Gossypium hirsutum and the origin of upland cotton. *American Journal of Botany*, 79, 1291 - 1310
- Wendel J.F., Schnabel A. & Seelanan T. (1995) Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (Gossypium). *Proceedings of the National Academy of Sciences*, 92, 280
- Wendel J.F. (2000) Genome evolution in polyploids. *Plant Molecular Biology*, 42, 225-249
- Whistler W.A.C. (1991) Polynesian plant introductions. In: *Islands, Plants, and Polynesians* (eds. Cox PA & Banack S), pp. 25-66. Dioscorides Press, Portland, Oregon
- White T.J., Bruns T., Lee S.B. & Taylor J.W. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols* (eds. Innis MA, Gelfand DH, Sninsky JJ & White TJ), pp. 315-322. Academic Press, San Diego
- Wiessner P. (2005) Social, symbolic, and ritual roles of the sweet potato in Enga, from its introduction until first contact. In: *The Sweet Potato in Oceania: a reappraisal*, pp. 121 - 130. Oceania Publications, Sydney
- Wills D.M. & Burke J.M. (2006) Chloroplast DNA Variation Confirms a Single Origin of Domesticated Sunflower (*Helianthus annuus* L.). *Journal of Heredity*, 97, 403-408
- Wilmshurst J.M., Anderson A.J., Higham T.F.G. & Worthy T.H. (2008) Dating the late prehistoric dispersal of polynesians to New Zealand using the commensal Pacific rat. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 7676-7680
- Wollstein A., Lao O., Becker C., Brauer S., Trent R.J., Nurnberg P., Stoneking M. & Kayser M. (2010) Demographic history of Oceania inferred from genome-wide data. *Curr Biol*, 20, 1983-92
- Wood D. & Lenne J.M. (1997) The conservation of agrobiodiversity on-farm: Questioning the emerging paradigm. *Biodiversity and Conservation*, 6, 109-129
- Wood T.E., Takebayashi N., Barker M.S., Mayrose I., Greenspoon P.B. & Rieseberg L.H. (2009) The frequency of polyploid speciation in vascular plants. *PNAS*, 106, 13875-13879
- Wright S.I., Bi I.V., Schroeder S.G., Yamasaki M., Doebley J.F., McMullen M.D. & Gaut B.S. (2005) The effects of artificial selection on the maize genome. *Science*, 308, 1310-4
- Xu D.H., Abe J., Gai J.Y. & Shimamoto Y. (2002) Diversity of chloroplast DNA SSRs in wild and cultivated soybeans: evidence for multiple origins of cultivated soybean. *Theoretical and Applied Genetics*, 105, 645-653

- Yen D.E. (1974) *The Sweet Potato in Oceania: An Essay in Ethnobotany*, Bishop Museum Press, Honolulu.
- Yen D.E. (1991) Domestication: The lesson from New Guinea. In: *Man and a half: Essays in Pacific anthropology and ethnobiology in honour of Ralph Bulmer* (ed. A P), pp. 558–569. Polynesian Society, Auckland
- Yen D.E. (1995) The development of Sahul agriculture with Australia as a bystander. *Antiquity*, 69, 831–847
- Yen D.E. (1998) Subsistence to commerce in Pacific agriculture: some four thousand years of plant exchange. In: *Plants for food and Medicine*. (eds. Pendergast HDV, Etkin NL, Houghton PJ & Harris DR). Royal Botanical Garden, Kew
- Zeder M.A., Emshwiller E., Smith B.D. & Bradley D.G. (2006a) Documenting domestication: the intersection of genetics and archaeology. *Trends in Genetics*, 22, 139-155
- Zeder M.a., Emshwiller E., Smith B.D. & Bradley D.G. (2006b) Documenting domestication: the intersection of genetics and archaeology. *Trends in Genetics*, 22, 139-55
- Zerega N.J., Ragone D. & Motley T.J. (2004) Complex origins of breadfruit (*Artocarpus altilis*, Moraceae): implications for human migrations in Oceania. *Am J Bot*, 91, 760-6
- Zhang D., Ghislain M., Huam Z., Golmirzaie A. & Hijmans R. (1998) RAPD variation in sweetpotato (*Ipomoea batatas* ( L .) Lam ) cultivars from South America and Papua New Guinea. *Genetic Resources Crop Evolution*, 45, 271-277
- Zhang D., Cervantes J., Huam Z., Carey E. & Ghislain M. (2000) Assessing genetic diversity of sweet potato (*Ipomoea batatas* ( L .) Lam .) cultivars from tropical America using AFLP. *Genetic Resources and Crop Evolution*, 659-665
- Zhang D., Rossel G., Kriegner A. & Hijmans R. (2004) AFLP assessment of diversity in sweetpotato from Latin America and the Pacific region: Its implications on the dispersal of the crop. *Genetic Resources and Crop Evolution*, 51, 115-120
- Zohary D. (2004) Unconscious selection and the evolution of domesticated plants. *Economic Botany*, 58, 5-10
- Zvelebil M. & Zvelebil K.V. (1988) Agricultural Transition and Indo-European Dispersals. *Antiquity*, 62, 574-583

## *Annexes*

**Annexe 1: Données passeport des échantillons analysés dans les différents chapitres de cette étude.**

a) Echantillons de *I. batatas* (Chapitres 3, 5 et 6)

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/ locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
1877	CTRAV	Vanuatu	Abe	Melanesia	Bankis woan manis	1_2	2	3	3	chap.6
1879	CTRAV	Vanuatu	Abe	Melanesia		1_2	2	16		chap.6
1878	CTRAV	Vanuatu	Abe	Melanesia	vouhi maréa	6_9	2	49		chap.6
1847	CTRAV	Vanuatu	Abe	Melanesia		1_2	2			chap.6
1891	CTRAV	Vanuatu	Abe	Melanesia	Daïtari	6_9	2			chap.6
1883	CTRAV	Vanuatu	Abe	Melanesia	Pumpkin	14	1			chap.6
1663	CTRAV	Vanuatu	Ambrym	Melanesia	Améré	1_2	2	15		chap.6
1606	CTRAV	Vanuatu	Ambrym	Melanesia	Améam	1_2	2	22		chap.6
1665	CTRAV	Vanuatu	Ambrym	Melanesia	Ral touto	6_9	2			chap.6
1661	CTRAV	Vanuatu	Ambrym	Melanesia	Areveka	14	1			chap.6
1660	CTRAV	Vanuatu	Ambrym	Melanesia	Ambokoou	ND	ND			chap.6
1607	CTRAV	Vanuatu	Ambrym	Melanesia	ahouo	6_9	2			chap.6
1664	CTRAV	Vanuatu	Ambrym	Melanesia	3 manis	1_2	2			chap.6
1666	CTRAV	Vanuatu	Ambrym	Melanesia	améam 2	ND	ND			chap.6
1668	CTRAV	Vanuatu	Ambrym	Melanesia	Tin pis	ND	ND			chap.6
1676	CTRAV	Vanuatu	Ambrym	Melanesia	Mélé	1_2	2			chap.6
Amb1	CTRAV	Vanuatu	Ambrym	Melanesia		1_2	2			chap.6
Amb2	CTRAV	Vanuatu	Ambrym	Melanesia		1_2	2			chap.6
Amb3	CTRAV	Vanuatu	Ambrym	Melanesia		1_2	2			chap.6
411	Yen	New Caledonia	Balade	Melanesia		1_2	2	1	1A	chap.6
410B	Yen	New Caledonia	Balade	Melanesia	kumala taabo	1_2	2	58	58	chap.6
410A	Yen	New Caledonia	Balade	Melanesia	kumala taabo	14	1	107		chap.6
415	Yen	New Caledonia	Balade	Melanesia		1_2	2			chap.6
410	Yen	New Caledonia	Balade	Melanesia	kumala taabo	1_2	2			chap.6
1959	CTRAV	Vanuatu	Banks	Melanesia	Sonat	1_2	2	3	3	chap.6
1953	CTRAV	Vanuatu	Banks	Melanesia	red head	ND	ND			chap.6
1946	CTRAV	Vanuatu	Banks	Melanesia	PNG	1_2	2			chap.6
banks1	CTRAV	Vanuatu	Banks	Melanesia		24	2			chap.6
217	LAES	The Bismarck Archipelago	Bougainville	Melanesia	Crisis	14	1	9		chap.6
23	LAES	The Bismarck Archipelago	Bougainville	Melanesia	Crisis	14	1	37		chap.6
24	LAES	The Bismarck Archipelago	Bougainville	Melanesia	NUG 2	6_9	2	40	40	chap.6
7	LAES	The Bismarck Archipelago	Bougainville	Melanesia	SIWAE	1_2	2			chap.6
8	LAES	The Bismarck Archipelago	Bougainville	Melanesia	WAK 3	1_2	2			chap.6
122	LAES	The Bismarck Archipelago	Bougainville	Melanesia	L 997	14	1			chap.6
151	LAES	The Bismarck Archipelago	Bougainville	Melanesia	NUG 5	14	1			chap.6
143	LAES	The Bismarck Archipelago	Bougainville	Melanesia	WAK 3	6_9	2			chap.6
21	LAES	The Bismarck Archipelago	Bougainville	Melanesia	NUG 5	14	1			chap.6
129	LAES	The Bismarck Archipelago	Bougainville	Melanesia	NUG 2	6_9	2			chap.6
211	LAES	The Bismarck Archipelago	Bougainville	Melanesia	Buka	1_2	2			chap.6



Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
44	LAES	The Bismarck Archipelago	Bougainville	Melanesia	L329	1_2	2			chap.6
49	LAES	The Bismarck Archipelago	Bougainville	Melanesia	L997	6_9	2			chap.6
400	Yen	New Caledonia	Chabwen	Melanesia	kumala miia rose	1_2	2	58	58	chap.6
404	Yen	New Caledonia	Chabwen	Melanesia	kumala belep	1_2	2	58	58	chap.6
10	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K9	1_2	2	4	4	chap.6
136	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	1_2	2	4	4	chap.6
180	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 37	1_2	2	4	4	chap.6
188	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	1_2	2	4	4	chap.6
45	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	1_2	2	4	4	chap.6
46	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	1_2	2	4	4	chap.6
4	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K9	1_2	2	4	4	chap.6
127	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 24	1_2	2	5	5	chap.6
116	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 36	14	1	9		chap.6
157	LAES	The Bismarck Archipelago	East New Britain	Melanesia	POI 13	1_2	2	24		chap.6
168	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	1_2	2	24		chap.6
158	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 37	14	1	25	25	chap.6
161	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 37	14	1	25	25	chap.6
167	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	14	1	25	25	chap.6
160	LAES	The Bismarck Archipelago	East New Britain	Melanesia	POI 13	1_2	2	26	26	chap.6
32	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 44	1_2	2	47		chap.6
52	LAES	The Bismarck Archipelago	East New Britain	Melanesia	L 135 Nomad 6	1_2	2	47		chap.6
177	LAES	The Bismarck Archipelago	East New Britain	Melanesia	POI 13	1_2	2			chap.6
163	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	1_2	2			chap.6
186	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	1_2	2			chap.6
175	LAES	The Bismarck Archipelago	East New Britain	Melanesia	POI 13	1_2	2			chap.6
90	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 54	1_2	2			chap.6
9	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K9	1_2	2			chap.6
47	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 43	14	1			chap.6
72	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K130 (L131)	6_9	2			chap.6
100	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 35	14	1			chap.6
169	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 37	1_2	2			chap.6
121	LAES	The Bismarck Archipelago	East New Britain	Melanesia	Doy 2	1_2	2			chap.6
126	LAES	The Bismarck Archipelago	East New Britain	Melanesia	Poi 6	1_2	2			chap.6
101	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 38	1_2	2			chap.6
115	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 45	14	1			chap.6
117	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	14	1			chap.6
118	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	1_2	2			chap.6
125	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 7	1_2	2			chap.6
128	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 142	6_9	2			chap.6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
135	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	6_9	2			chap.6
138	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 7	1_2	2			chap.6
14	LAES	The Bismarck Archipelago	East New Britain	Melanesia	Doy 2	14	1			chap.6
159	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	1_2	2			chap.6
174	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 37	14	1			chap.6
176	LAES	The Bismarck Archipelago	East New Britain	Melanesia	POI 13	1_2	2			chap.6
17	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 142	6_9	2			chap.6
187	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	14	1			chap.6
33	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 45	14	1			chap.6
41	LAES	The Bismarck Archipelago	East New Britain	Melanesia	POI 6	14	1			chap.6
54	LAES	The Bismarck Archipelago	East New Britain	Melanesia	RAB 24	14	1			chap.6
5	LAES	The Bismarck Archipelago	East New Britain	Melanesia	POI 13	1_2	2			chap.6
153	LAES	The Bismarck Archipelago	East New Britain	Melanesia	K 9	1_2	2			chap.6
361	Yen	The Bismarck Archipelago	East New Britain, Rabaul	Melanesia	buka	14	1			chap.6
358	Yen	The Bismarck Archipelago	East New Britain, Rabaul	Melanesia	tapuka	1_2	2			chap.6
362	Yen	The Bismarck Archipelago	East New Britain, Rabaul	Melanesia	kambubu	1_2	2			chap.6
1531	CTRAV	Vanuatu	Epi	Melanesia	imala bom à feuilles pointu	1_2	2	22		chap.6
1532	CTRAV	Vanuatu	Epi	Melanesia	Blancviolet	1_2	2	23		chap.6
1568	CTRAV	Vanuatu	Epi	Melanesia	tongoa	1_2	2	23		chap.6
1359	CTRAV	Vanuatu	Eromango	Melanesia	Meliang	1_2	2	15		chap.6
1361	CTRAV	Vanuatu	Eromango	Melanesia	ndan navlar	1_2	2	16		chap.6
1360	CTRAV	Vanuatu	Eromango	Melanesia	nosso viliao	30	1			chap.6
1397	CTRAV	Vanuatu	Eromango	Melanesia	Tiaré	28	2			chap.6
1396	CTRAV	Vanuatu	Eromango	Melanesia	Noklintari	ND	ND			chap.6
367	Yen	Solomon Islands	Malaita	Melanesia	sangobolu	1_2	2			chap.6
368	Yen	Solomon Islands	Malaita	Melanesia		1_2	2			chap.6
1432	CTRAV	Vanuatu	Malekula	Melanesia	Blanche 2	1_2	2	16		chap.6
1433	CTRAV	Vanuatu	Malekula	Melanesia	Blanche 3	1_2	2			chap.6
1431	CTRAV	Vanuatu	Malekula	Melanesia	blanche 1	14	1			chap.6
1064	CTRAV	Vanuatu	Malo	Melanesia	Wasuru suru hao	14	1	2		chap.6
146	LAES	The Bismarck Archipelago	Manus	Melanesia	MAS 1	1_2	2			chap.6
221	LAES	The Bismarck Archipelago	Manus	Melanesia	Nauna Is	1_2	2			chap.6
225	LAES	The Bismarck Archipelago	Manus	Melanesia	LP 01 Mokereng vill	1_2	2			chap.6
35	LAES	The Bismarck Archipelago	Manus	Melanesia	MAS 1	1_2	2			chap.6
545_SOL	Yen	Solomon Islands	Moorea	Melanesia		1_2	2			chap.6
166	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 61	1_2	2	4	4	chap.6
131	LAES	The Bismarck Archipelago	New Ireland	Melanesia	L 942 Kavieng 59	14	1	14	14	chap.6
164	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 79	14	1	25	25	chap.6
165	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 79	1_2	2	26	26	chap.6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
155	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 79	1_2	2			chap.6
89	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 57	1_2	2			chap.6
170	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 79	1_2	2			chap.6
179	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 79	1_2	2			chap.6
162	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 61	1_2	2			chap.6
85	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 99	1_2	2			chap.6
114	LAES	The Bismarck Archipelago	New Ireland	Melanesia	L 949 Kavieng 90	1_2	2			chap.6
149	LAES	The Bismarck Archipelago	New Ireland	Melanesia	L 949 Kavieng 90	1_2	2			chap.6
12	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 11	1_2	2			chap.6
156	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 61	1_2	2			chap.6
178	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 61	1_2	2			chap.6
145	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 61	6_9	2			chap.6
148	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 11	6_9	2			chap.6
20	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 81	1_2	2			chap.6
26	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 34	6_9	2			chap.6
37	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 79	1_2	2			chap.6
59	LAES	The Bismarck Archipelago	New Ireland	Melanesia	L 942 Kavieng 59	1_2	2			chap.6
93	LAES	The Bismarck Archipelago	New Ireland	Melanesia	KAV 57	14	1			chap.6
1770	CTRAV	Vanuatu	Pentecost	Melanesia	komal rubap	1_2	2	95	95	chap.6
1793	CTRAV	Vanuatu	Pentecost	Melanesia	Saout	1_2	2	95	95	chap.6
1801	CTRAV	Vanuatu	Pentecost	Melanesia	Kumala 4	30	1			chap.6
1802	CTRAV	Vanuatu	Pentecost	Melanesia	Kumala 5	6_9	2			chap.6
1745	CTRAV	Vanuatu	Pentecost	Melanesia	Kumala 2	14	1			chap.6
1792	CTRAV	Vanuatu	Pentecost	Melanesia	kumala 3	6_9	2			chap.6
1148	CTRAV	Vanuatu	Santo	Melanesia	Kumala Takar	ND	ND	7		chap.6
1165	CTRAV	Vanuatu	Santo	Melanesia	peavut	1_2	2	8	8	chap.6
1164	CTRAV	Vanuatu	Santo	Melanesia	Latsmoli	6_9	2	8	8	chap.6
1167	CTRAV	Vanuatu	Santo	Melanesia	Kumala Anna	14	1			chap.6
1162	CTRAV	Vanuatu	Santo	Melanesia	Lomaeto	ND	ND			chap.6
1149	CTRAV	Vanuatu	Santo	Melanesia		ND	ND			chap.6
1163	CTRAV	Vanuatu	Santo	Melanesia	Bernadette	14	1			chap.6
1166	CTRAV	Vanuatu	Santo	Melanesia	Paokin	1_2	2			chap.6
382	Yen	Vanuatu	Santo	Melanesia		6_9	2	49		chap.6
379	Yen	Vanuatu	Santo	Melanesia		14	1	50	50	chap.6
381	Yen	Vanuatu	Santo	Melanesia	bina/sensi	1_2	2	52	52	chap.6
386	Yen	Vanuatu	Santo	Melanesia		14	1			chap.6
375	Yen	Vanuatu	Santo	Melanesia		14	1			chap.6
376	Yen	Vanuatu	Santo	Melanesia		1_2	2			chap.6
377	Yen	Vanuatu	Santo	Melanesia		1_2	2			chap.6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
1306	CTRAV	Vanuatu	Tanna	Melanesia		28	2	7		chap.6
1292	CTRAV	Vanuatu	Tanna	Melanesia	Iodile	6_9	2			chap.6
1305	CTRAV	Vanuatu	Tanna	Melanesia	agrikalja	1_2	2			chap.6
1304	CTRAV	Vanuatu	Tanna	Melanesia	Aope	ND	ND			chap.6
1307	CTRAV	Vanuatu	Tanna	Melanesia	Koleilbwouo	1_2	2			chap.6
1318	CTRAV	Vanuatu	Tanna	Melanesia	potete	14	1			chap.6
397	Yen	Vanuatu	Vila	Melanesia		14	1	52	52	chap.6
398	Yen	Vanuatu	Vila	Melanesia		1_2	2			chap.6
389	Yen	Vanuatu	West Santo	Melanesia		1_2	2	52	52	chap.6
387	Yen	Vanuatu	West Santo	Melanesia		14	1			chap.6
440399	CIP	New Caledonia		Melanesia		14	1			chap.6
440202	CIP	Solomon Islands		Melanesia		1_2	2			chap.6
440360	CIP	Solomon Islands		Melanesia		1_2	2			chap.6
440362	CIP	Solomon Islands		Melanesia		14	1			chap.6
440363	CIP	Solomon Islands		Melanesia		1_2	2			chap.6
CTRAV_1	CTRAV	Vanuatu		Melanesia		1_2	2			chap.6
ply20	CTRAV	Vanuatu		Melanesia		29	1			chap.6
rv30	CTRAV	Vanuatu		Melanesia		1_2	2			chap.6
salilli	CTRAV	Vanuatu		Melanesia	Salili	27	2			chap.6
141	LAES	Solomon Islands		Melanesia		1_2	2	18		chap.6
181	LAES	Solomon Islands		Melanesia		14	1	25	25	chap.6
25	LAES	Solomon Islands		Melanesia		14	1	37	37	chap.6
66	LAES	Solomon Islands		Melanesia		ND	ND	37	37	chap.6
48	LAES	Solomon Islands		Melanesia		ND	ND	40	40	chap.6
55	LAES	Solomon Islands		Melanesia		6_9	2	40	40	chap.6
183	LAES	Solomon Islands		Melanesia		6_9	2	96	96	chap.6
192	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
74	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
197	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
34	LAES	Solomon Islands		Melanesia		ND	ND			chap.6
82	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
193	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
22	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
36	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
57	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
84	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
119	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
18	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
86	LAES	Solomon Islands		Melanesia		1_2	2			chap.6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
92	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
120	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
13	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
38	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
39	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
40	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
60	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
65	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
67	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
6	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
98	LAES	Solomon Islands		Melanesia		1_2	2			chap.6
110	LAES	Vanuatu		Melanesia		6_9	2	5	5	chap.6
130	LAES	Vanuatu		Melanesia		14	1	14	14	chap.6
95	LAES	Vanuatu		Melanesia		6_9	2	96	96	chap.6
79	LAES	Vanuatu		Melanesia		14	1			chap.6
69	LAES	Vanuatu		Melanesia		6_9	2			chap.6
182	LAES	Vanuatu		Melanesia		1_2	2			chap.6
96	LAES	Vanuatu		Melanesia		1_2	2			chap.6
427	Yen	Fiji		Melanesia	samoa	1_2	2	1	1A	chap.6
425	Yen	Fiji		Melanesia	red timala	14	1	50	50	chap.6
428	Yen	Fiji		Melanesia	V52	1_2	2	66		chap.6
426	Yen	Fiji		Melanesia	white timala	14	1	107		chap.6
429	Yen	Fiji		Melanesia	navuso	14	1			chap.6
432	Yen	Fiji		Melanesia	cocolevu	14	1			chap.6
231	Yen	West New Guinea	Baliem	New Guinea		14	1	36		chap.5 et 6
230	Yen	West New Guinea	Baliem	New Guinea	helemaiu	14	1			chap.5 et 6
232	Yen	West New Guinea	Baliem	New Guinea	hulok	1_2	2			chap.5 et 6
233	Yen	West New Guinea	Baliem	New Guinea	werene	1_2	2			chap.5 et 6
234	Yen	West New Guinea	Baliem	New Guinea	hiwesa	1_2	2			chap.5 et 6
235	Yen	West New Guinea	Baliem	New Guinea	lyoka	1_2	2			chap.5 et 6
240_1	Yen	West New Guinea	Biak	New Guinea	ansio berik	1_2	2			chap.5 et 6
240	Yen	West New Guinea	Biak	New Guinea	ansio berik	ND	ND			chap.5 et 6
240_2	Yen	West New Guinea	Biak	New Guinea	ansio berik	1_2	2			chap.5 et 6
68	LAES	Papua New Guinea	Central	New Guinea	L 188 Godugudu	6_9	2			chap.5 et 6
201	LAES	Papua New Guinea	East Sepik	New Guinea	Yar	14	1			chap.5 et 6
973	HAES	Papua New Guinea	Eastern Highland	New Guinea	Kurina	1_2	2	6		chap.5 et 6
639	HAES	Papua New Guinea	Eastern Highland	New Guinea		1_2	2	65	65	chap.5 et 6
559	HAES	Papua New Guinea	Eastern Highland	New Guinea	Munduenta	1_2	2	69		chap.5 et 6
624	HAES	Papua New Guinea	Eastern Highland	New Guinea	Aiden?	1_2	2	69		chap.5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
630	HAES	Papua New Guinea	Eastern Highland	New Guinea		13	1	71		chap.5 et 6
643	HAES	Papua New Guinea	Eastern Highland	New Guinea	Kagip?	14	1	90	90	chap.5 et 6
636	HAES	Papua New Guinea	Eastern Highland	New Guinea	Aiden	14	1	92	92	chap.5 et 6
626	HAES	Papua New Guinea	Eastern Highland	New Guinea	Serenta Tokame?	6_9	2			chap.5 et 6
4	HAES	Papua New Guinea	Eastern Highland	New Guinea		1_2	2			chap.5 et 6
10	HAES	Papua New Guinea	Eastern Highland	New Guinea		6_9	2			chap.5 et 6
563	HAES	Papua New Guinea	Eastern Highland	New Guinea	Kekori	14	1			chap.5 et 6
628	HAES	Papua New Guinea	Eastern Highland	New Guinea	Wanmun Red	14	1			chap.5 et 6
629	HAES	Papua New Guinea	Eastern Highland	New Guinea		1_2	2			chap.5 et 6
640	HAES	Papua New Guinea	Eastern Highland	New Guinea	Sekamapa?	6_9	2			chap.5 et 6
641	HAES	Papua New Guinea	Eastern Highland	New Guinea	Chekel?	14	1			chap.5 et 6
6	HAES	Papua New Guinea	Eastern Highland	New Guinea	Merikan	1_2	2			chap.5 et 6
77	LAES	Papua New Guinea	Eastern Highland	New Guinea	Kekori	6_9	2			chap.5 et 6
275	Yen	Papua New Guinea	Eastern Highlands, Aiyura	New Guinea	(serenta)	14	1	36		chap.5 et 6
277	Yen	Papua New Guinea	Eastern Highlands, Namura	New Guinea	hagesya	1_2	2	41		chap.5 et 6
276	Yen	Papua New Guinea	Eastern Highlands, Namura	New Guinea	sigarup	1_2	2			chap.5 et 6
281	Yen	Papua New Guinea	Eastern Highlands, Namura	New Guinea	yuka	14	1			chap.5 et 6
282	Yen	Papua New Guinea	Eastern Highlands, Namura	New Guinea	afaioea	14	1			chap.5 et 6
283	Yen	Papua New Guinea	Eastern Highlands, Namura	New Guinea	mamat	1_2	2			chap.5 et 6
267	Yen	Papua New Guinea	Eastern Highlands, Watabung	New Guinea	foravena	1_2	2			chap.5 et 6
269	Yen	Papua New Guinea	Eastern Highlands, Watabung	New Guinea	fatena	1_2	2			chap.5 et 6
270	Yen	Papua New Guinea	Eastern Highlands, Watabung	New Guinea	obume	14	1			chap.5 et 6
271	Yen	Papua New Guinea	Eastern Highlands, Watabung	New Guinea	kiravena	1_2	2			chap.5 et 6
515	HAES	Papua New Guinea	Gulf	New Guinea		1_2	2	31	31	chap.5 et 6
399	HAES	Papua New Guinea	Gulf	New Guinea	Kalarumeasea	1_2	2	44	44	chap.5 et 6
516	HAES	Papua New Guinea	Gulf	New Guinea	Popale	14	1	97	97	chap.5 et 6
70	LAES	Papua New Guinea	Gulf	New Guinea	L 22 Murua Wanmun 1	1_2	2			chap.5 et 6
219	Yen	West New Guinea	Hollandia	New Guinea	menes poetih	14	1			chap.5 et 6
401	HAES	Papua New Guinea	Madang	New Guinea	Wanmun Madang	12	1			chap.5 et 6
83	LAES	Papua New Guinea	Madang	New Guinea	L 312	1_2	2			chap.5 et 6
144	LAES	Papua New Guinea	Madang	New Guinea	L 329 Kumbani	1_2	2			chap.5 et 6
185	LAES	Papua New Guinea	Madang	New Guinea	L 312	1_2	2			chap.5 et 6
94	LAES	Papua New Guinea	Madang	New Guinea	L 318	14	1			chap.5 et 6
214_1	Yen	West New Guinea	Merauke	New Guinea	notta waikio	14	1	34		chap.5 et 6
213	Yen	West New Guinea	Merauke	New Guinea	notta waikio	14	1			chap.5 et 6
214_2	Yen	West New Guinea	Merauke	New Guinea	notta waikio	14	1			chap.5 et 6
215	Yen	West New Guinea	Merauke	New Guinea	notta waikio	14	1			chap.5 et 6
217	Yen	West New Guinea	Merauke	New Guinea	notta waikio	14	1			chap.5 et 6
619	HAES	Papua New Guinea	Milne Bay	New Guinea	Toiaw	1_2	2	30		chap.5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
18	HAES	Papua New Guinea	Milne Bay	New Guinea	Gauguta	1_2	2	31	31	chap.5 et 6
668	HAES	Papua New Guinea	Milne Bay	New Guinea	Oscan	1_2	2	46		chap.5 et 6
420	HAES	Papua New Guinea	Milne Bay	New Guinea	Castaboga	1_2	2	65		chap.5 et 6
671	HAES	Papua New Guinea	Milne Bay	New Guinea	Semani	13	1	71		chap.5 et 6
620	HAES	Papua New Guinea	Milne Bay	New Guinea	Kaintabara	1_2	2	94	94	chap.5 et 6
664	HAES	Papua New Guinea	Milne Bay	New Guinea	Ite	1_2	2	94	94	chap.5 et 6
600	HAES	Papua New Guinea	Milne Bay	New Guinea	Paiwa	1_2	2			chap.5 et 6
71	HAES	Papua New Guinea	Milne Bay	New Guinea	Taie	1_2	2			chap.5 et 6
423	HAES	Papua New Guinea	Milne Bay	New Guinea	Gorasip	13	1			chap.5 et 6
421	HAES	Papua New Guinea	Milne Bay	New Guinea	Ipau	1_2	2			chap.5 et 6
467	HAES	Papua New Guinea	Milne Bay	New Guinea		6_9	2			chap.5 et 6
162	HAES	Papua New Guinea	Milne Bay	New Guinea	Wanmun	14	1			chap.5 et 6
164	HAES	Papua New Guinea	Milne Bay	New Guinea	Bouyavani	1_2	2			chap.5 et 6
165	HAES	Papua New Guinea	Milne Bay	New Guinea	Taivit	14	1			chap.5 et 6
17	HAES	Papua New Guinea	Milne Bay	New Guinea	Makimar	6_9	2			chap.5 et 6
38	HAES	Papua New Guinea	Milne Bay	New Guinea	Kubol	1_2	2			chap.5 et 6
40	HAES	Papua New Guinea	Milne Bay	New Guinea		6_9	2			chap.5 et 6
422	HAES	Papua New Guinea	Milne Bay	New Guinea	Popondetta	14	1			chap.5 et 6
424	HAES	Papua New Guinea	Milne Bay	New Guinea	yavanpom	14	1			chap.5 et 6
43	HAES	Papua New Guinea	Milne Bay	New Guinea	Waip	1_2	2			chap.5 et 6
672	HAES	Papua New Guinea	Milne Bay	New Guinea	Kankanmun	14	1			chap.5 et 6
602	HAES	Papua New Guinea	Milne Bay	New Guinea	Dibura	1_2	2			chap.5 et 6
109	LAES	Papua New Guinea	Milne Bay	New Guinea	L 46 Alotau	1_2	2	3	3	chap.5 et 6
69	HAES	Papua New Guinea	Morobe	New Guinea		1_2	2	43	43	chap.5 et 6
911	HAES	Papua New Guinea	Morobe	New Guinea	Wantoot	1_2	2	48		chap.5 et 6
872	HAES	Papua New Guinea	Morobe	New Guinea	Romba	14	1			chap.5 et 6
182	HAES	Papua New Guinea	Morobe	New Guinea	Wasamia	14	1			chap.5 et 6
183	HAES	Papua New Guinea	Morobe	New Guinea	Timini	14	1			chap.5 et 6
1003	HAES	Papua New Guinea	Morobe	New Guinea	Mare	1_2	2			chap.5 et 6
676	HAES	Papua New Guinea	Morobe	New Guinea	Chimbu	14	1			chap.5 et 6
893	HAES	Papua New Guinea	Morobe	New Guinea	Aruguay	6_9	2			chap.5 et 6
907	HAES	Papua New Guinea	Morobe	New Guinea	Lae 2	14	1			chap.5 et 6
19	LAES	Papua New Guinea	Morobe	New Guinea	B 11	1_2	2			chap.5 et 6
137	LAES	Papua New Guinea	Morobe	New Guinea	B 11	1_2	2			chap.5 et 6
147	LAES	Papua New Guinea	Morobe	New Guinea	BUB 1	1_2	2			chap.5 et 6
42	LAES	Papua New Guinea	Morobe	New Guinea	BUB 1	1_2	2			chap.5 et 6
175	HAES	Papua New Guinea	Oro	New Guinea		14	1			chap.5 et 6
491	HAES	Papua New Guinea	Oro	New Guinea		14	1			chap.5 et 6
174	HAES	Papua New Guinea	Oro	New Guinea		1_2	2			chap.5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
64	LAES	Papua New Guinea	Oro	New Guinea	L 9 Imi du Gogie	1_2	2	4	4	chap.5 et 6
142	LAES	Papua New Guinea	Oro	New Guinea	L 676	1_2	2	18		chap.5 et 6
62	LAES	Papua New Guinea	Oro	New Guinea	L 676	6_9	2			chap.5 et 6
202	LAES	Papua New Guinea	Oro	New Guinea	Kaura	1_2	2			chap.5 et 6
61	LAES	Papua New Guinea	Oro	New Guinea	L 43	14	1			chap.5 et 6
132	LAES	Papua New Guinea	Oro	New Guinea	L 879	14	1			chap.5 et 6
152	LAES	Papua New Guinea	Oro	New Guinea	L 9	14	1			chap.5 et 6
31	LAES	Papua New Guinea	Oro	New Guinea	L 879	1_2	2			chap.5 et 6
253	Yen	Papua New Guinea	PT Moresby	New Guinea		1_2	2	41		chap.5 et 6
257	Yen	Papua New Guinea	PT Moresby	New Guinea		1_2	2	41		chap.5 et 6
246	Yen	Papua New Guinea	PT Moresby	New Guinea		1_2	2			chap.5 et 6
2	HAES	Papua New Guinea	Sandaun	New Guinea	Mea	6_9	2	20	20	chap.5 et 6
550	HAES	Papua New Guinea	Sandaun	New Guinea		1_2	2	30	30	chap.5 et 6
552	HAES	Papua New Guinea	Sandaun	New Guinea		1_2	2			chap.5 et 6
843	HAES	Papua New Guinea	Sandaun	New Guinea		1_2	2			chap.5 et 6
558	HAES	Papua New Guinea	Sandaun	New Guinea		14	1			chap.5 et 6
549	HAES	Papua New Guinea	Sandaun	New Guinea		1_2	2			chap.5 et 6
551	HAES	Papua New Guinea	Sandaun	New Guinea		1_2	2			chap.5 et 6
16	LAES	Papua New Guinea	Sandaun	New Guinea	PNG 692	1_2	2			chap.5 et 6
133	LAES	Papua New Guinea	Sandaun	New Guinea	PNG 692	14	1			chap.5 et 6
118	HAES	Papua New Guinea	Simbu	New Guinea	Amp Bilkambu	14	1	10		chap.5 et 6
129	HAES	Papua New Guinea	Simbu	New Guinea	Muno Gudunga	14	1	10	10	chap.5 et 6
130	HAES	Papua New Guinea	Simbu	New Guinea	Taindi Ambu	1_2	2	13		chap.5 et 6
30	HAES	Papua New Guinea	Simbu	New Guinea	Krukanambu	1_2	2	13	13	chap.5 et 6
142	HAES	Papua New Guinea	Simbu	New Guinea		6_9	2	20	20	chap.5 et 6
19	HAES	Papua New Guinea	Simbu	New Guinea	Galimbol	6_9	2	20	20	chap.5 et 6
52	HAES	Papua New Guinea	Simbu	New Guinea		6_9	2	20	20	chap.5 et 6
150	HAES	Papua New Guinea	Simbu	New Guinea		6_9	2	21		chap.5 et 6
201	HAES	Papua New Guinea	Simbu	New Guinea	Bomai Kruo	1_2	2	29		chap.5 et 6
75	HAES	Papua New Guinea	Simbu	New Guinea		1_2	2	31		chap.5 et 6
80	HAES	Papua New Guinea	Simbu	New Guinea		1_2	2	43	43	chap.5 et 6
79	HAES	Papua New Guinea	Simbu	New Guinea		1_2	2	43		chap.5 et 6
365	HAES	Papua New Guinea	Simbu	New Guinea	Warri	1_2	2	44	44	chap.5 et 6
823	HAES	Papua New Guinea	Simbu	New Guinea	Dowaggi	1_2	2	44	44	chap.5 et 6
831	HAES	Papua New Guinea	Simbu	New Guinea		14	1	80		chap.5 et 6
803	HAES	Papua New Guinea	Simbu	New Guinea		14	1	81	81	chap.5 et 6
93	HAES	Papua New Guinea	Simbu	New Guinea	Mondambo	14	1	81	81	chap.5 et 6
804	HAES	Papua New Guinea	Simbu	New Guinea		14	1	82		chap.5 et 6
806	HAES	Papua New Guinea	Simbu	New Guinea		14	1	82		chap.5 et 6



Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
202	HAES	Papua New Guinea	Simbu	New Guinea	Wanmun yellow	1_2	2	89	89	chap.5 et 6
356	HAES	Papua New Guinea	Simbu	New Guinea	Konomanggi	14	1	97	97	chap.5 et 6
145	HAES	Papua New Guinea	Simbu	New Guinea		13	1			chap.5 et 6
523	HAES	Papua New Guinea	Simbu	New Guinea	Wanmun	13	1			chap.5 et 6
415	HAES	Papua New Guinea	Simbu	New Guinea		14	1			chap.5 et 6
820	HAES	Papua New Guinea	Simbu	New Guinea	Kaukau	14	1			chap.5 et 6
144	HAES	Papua New Guinea	Simbu	New Guinea		1_2	2			chap.5 et 6
149	HAES	Papua New Guinea	Simbu	New Guinea		6_9	2			chap.5 et 6
511	HAES	Papua New Guinea	Simbu	New Guinea		14	1			chap.5 et 6
532	HAES	Papua New Guinea	Simbu	New Guinea	Daccambo	6_9	2			chap.5 et 6
138	HAES	Papua New Guinea	Simbu	New Guinea		1_2	2			chap.5 et 6
816	HAES	Papua New Guinea	Simbu	New Guinea	Kruabal	1_2	2			chap.5 et 6
814	HAES	Papua New Guinea	Simbu	New Guinea	Urauranggi	6_9	2			chap.5 et 6
133	HAES	Papua New Guinea	Simbu	New Guinea	Mimbik Kamanamb	13	1			chap.5 et 6
528	HAES	Papua New Guinea	Simbu	New Guinea	Ogan Gaigol	14	1			chap.5 et 6
34	HAES	Papua New Guinea	Simbu	New Guinea	Diniambo	1_2	2			chap.5 et 6
91	HAES	Papua New Guinea	Simbu	New Guinea		1_2	2			chap.5 et 6
160	HAES	Papua New Guinea	Simbu	New Guinea		14	1			chap.5 et 6
538	HAES	Papua New Guinea	Simbu	New Guinea	Parakuambo	14	1			chap.5 et 6
599	HAES	Papua New Guinea	Simbu	New Guinea	Sainamb White	14	1			chap.5 et 6
358	HAES	Papua New Guinea	Simbu	New Guinea	Angamolabal	14	1			chap.5 et 6
366	HAES	Papua New Guinea	Simbu	New Guinea	Gerrowaggi	1_2	2			chap.5 et 6
439	HAES	Papua New Guinea	Simbu	New Guinea		14	1			chap.5 et 6
535	HAES	Papua New Guinea	Simbu	New Guinea	From Ramu	1_2	2			chap.5 et 6
203	HAES	Papua New Guinea	Simbu	New Guinea	Kamambu	14	1			chap.5 et 6
443	HAES	Papua New Guinea	Simbu	New Guinea	Morrogi	6_9	2			chap.5 et 6
134	HAES	Papua New Guinea	Simbu	New Guinea	Indaun Mulgoamb	6_9	2			chap.5 et 6
139	HAES	Papua New Guinea	Simbu	New Guinea	Nilkugu	1_2	2			chap.5 et 6
155	HAES	Papua New Guinea	Simbu	New Guinea		14	1			chap.5 et 6
199	HAES	Papua New Guinea	Simbu	New Guinea	Duininamb	1_2	2			chap.5 et 6
200	HAES	Papua New Guinea	Simbu	New Guinea	Kumambu	1_2	2			chap.5 et 6
20	HAES	Papua New Guinea	Simbu	New Guinea	Bomai	1_2	2			chap.5 et 6
33	HAES	Papua New Guinea	Simbu	New Guinea	Maroabal	6_9	2			chap.5 et 6
357	HAES	Papua New Guinea	Simbu	New Guinea	Moromaggi	14	1			chap.5 et 6
36	HAES	Papua New Guinea	Simbu	New Guinea	Planggi Demambo	1_2	2			chap.5 et 6
37	HAES	Papua New Guinea	Simbu	New Guinea	Ambgamb	14	1			chap.5 et 6
442	HAES	Papua New Guinea	Simbu	New Guinea	Tripelanggi	14	1			chap.5 et 6
444	HAES	Papua New Guinea	Simbu	New Guinea	Kaniggi	14	1			chap.5 et 6
512	HAES	Papua New Guinea	Simbu	New Guinea		14	1			chap.5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
520	HAES	Papua New Guinea	Simbu	New Guinea		13	1			chap.5 et 6
527	HAES	Papua New Guinea	Simbu	New Guinea	Galu Gai	1_2	2			chap.5 et 6
529	HAES	Papua New Guinea	Simbu	New Guinea	Makmakgai	14	1			chap.5 et 6
534	HAES	Papua New Guinea	Simbu	New Guinea	Gaiambo	1_2	2			chap.5 et 6
536	HAES	Papua New Guinea	Simbu	New Guinea	Yomboambu	6_9	2			chap.5 et 6
537	HAES	Papua New Guinea	Simbu	New Guinea	Taindambo	1_2	2			chap.5 et 6
541	HAES	Papua New Guinea	Simbu	New Guinea	Simbugambo	1_2	2			chap.5 et 6
598	HAES	Papua New Guinea	Simbu	New Guinea	Sainamb Red	1_2	2			chap.5 et 6
608	HAES	Papua New Guinea	Simbu	New Guinea		1_2	2			chap.5 et 6
807	HAES	Papua New Guinea	Simbu	New Guinea		14	1			chap.5 et 6
808	HAES	Papua New Guinea	Simbu	New Guinea		6_9	2			chap.5 et 6
812	HAES	Papua New Guinea	Simbu	New Guinea	Nondugimam	1_2	2			chap.5 et 6
817	HAES	Papua New Guinea	Simbu	New Guinea	Wanmun red	14	1			chap.5 et 6
822	HAES	Papua New Guinea	Simbu	New Guinea	Boggogig	1_2	2			chap.5 et 6
824	HAES	Papua New Guinea	Simbu	New Guinea	Sikikruo	1_2	2			chap.5 et 6
113	LAES	Papua New Guinea	Simbu	New Guinea	PNG 1164	1_2	2	6		chap.5 et 6
186	HAES	Papua New Guinea	Southern Highland	New Guinea	Koneme	14	1	10	10	chap.5 et 6
216	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	14	1	10	10	chap.5 et 6
437	HAES	Papua New Guinea	Southern Highland	New Guinea	Ma	1_2	2	13	13	chap.5 et 6
205	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	6_9	2	21		chap.5 et 6
189	HAES	Papua New Guinea	Southern Highland	New Guinea	Popenongh	1_2	2	30	30	chap.5 et 6
299	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	1_2	2	30	30	chap.5 et 6
747	HAES	Papua New Guinea	Southern Highland	New Guinea	Popo	1_2	2	30		chap.5 et 6
326	HAES	Papua New Guinea	Southern Highland	New Guinea		1_2	2	31	31	chap.5 et 6
332	HAES	Papua New Guinea	Southern Highland	New Guinea		1_2	2	31	31	chap.5 et 6
735	HAES	Papua New Guinea	Southern Highland	New Guinea	Pigawe	1_2	2	31	31	chap.5 et 6
246	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	14	1	38	38	chap.5 et 6
252	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	14	1	38	38	chap.5 et 6
272	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	1_2	2	42		chap.5 et 6
380	HAES	Papua New Guinea	Southern Highland	New Guinea	Lumbawa	1_2	2	42		chap.5 et 6
27	HAES	Papua New Guinea	Southern Highland	New Guinea	Aiyuku Ina	1_2	2	43	43	chap.5 et 6
300	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	1_2	2	44	44	chap.5 et 6
310	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	1_2	2	44	44	chap.5 et 6
354	HAES	Papua New Guinea	Southern Highland	New Guinea	K1 (Breeding)	1_2	2	44	44	chap.5 et 6
749	HAES	Papua New Guinea	Southern Highland	New Guinea	Aron	1_2	2	44	44	chap.5 et 6
309	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	14	1	45	45	chap.5 et 6
355	HAES	Papua New Guinea	Southern Highland	New Guinea	K2 (Breeding)	14	1	45	45	chap.5 et 6
319	HAES	Papua New Guinea	Southern Highland	New Guinea		1_2	2	46		chap.5 et 6
334	HAES	Papua New Guinea	Southern Highland	New Guinea		1_2	2	48		chap.5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
734	HAES	Papua New Guinea	Southern Highland	New Guinea	Habia	1_2	2	65	65	chap.5 et 6
712	HAES	Papua New Guinea	Southern Highland	New Guinea		1_2	2	76		chap.5 et 6
753	HAES	Papua New Guinea	Southern Highland	New Guinea	Koe	1_2	2	76		chap.5 et 6
800	HAES	Papua New Guinea	Southern Highland	New Guinea	Gurume Gai	14	1	80		chap.5 et 6
188	HAES	Papua New Guinea	Southern Highland	New Guinea		1_2	2	89	89	chap.5 et 6
497	HAES	Papua New Guinea	Southern Highland	New Guinea	Kake Ina	14	1	90	90	chap.5 et 6
710	HAES	Papua New Guinea	Southern Highland	New Guinea		14	1	90	90	chap.5 et 6
711	HAES	Papua New Guinea	Southern Highland	New Guinea		14	1	92	92	chap.5 et 6
718	HAES	Papua New Guinea	Southern Highland	New Guinea		14	1	92	92	chap.5 et 6
311	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	1_2	2	94	94	chap.5 et 6
748	HAES	Papua New Guinea	Southern Highland	New Guinea	Lain	14	1	97	97	chap.5 et 6
242	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	6_9	2			chap.5 et 6
353	HAES	Papua New Guinea	Southern Highland	New Guinea	K14 (Ex Mamme)	12	1			chap.5 et 6
746	HAES	Papua New Guinea	Southern Highland	New Guinea	Pori	13	1			chap.5 et 6
342	HAES	Papua New Guinea	Southern Highland	New Guinea		14	1			chap.5 et 6
344	HAES	Papua New Guinea	Southern Highland	New Guinea	Naveto	14	1			chap.5 et 6
772	HAES	Papua New Guinea	Southern Highland	New Guinea	Pulupirri	13	1			chap.5 et 6
362	HAES	Papua New Guinea	Southern Highland	New Guinea	Kariap hake	12	1			chap.5 et 6
243	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	14	1			chap.5 et 6
320	HAES	Papua New Guinea	Southern Highland	New Guinea		14	1			chap.5 et 6
303	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	14	1			chap.5 et 6
317	HAES	Papua New Guinea	Southern Highland	New Guinea	Sapel	14	1			chap.5 et 6
264	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	14	1			chap.5 et 6
101	HAES	Papua New Guinea	Southern Highland	New Guinea	Lagapin	6_9	2			chap.5 et 6
740	HAES	Papua New Guinea	Southern Highland	New Guinea	Mendkala	6_9	2			chap.5 et 6
247	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	1_2	2			chap.5 et 6
298	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	1_2	2			chap.5 et 6
503	HAES	Papua New Guinea	Southern Highland	New Guinea	Kove	13	1			chap.5 et 6
514	HAES	Papua New Guinea	Southern Highland	New Guinea	Hono Banana	1_2	2			chap.5 et 6
100	HAES	Papua New Guinea	Southern Highland	New Guinea	Hagenaik	6_9	2			chap.5 et 6
169	HAES	Papua New Guinea	Southern Highland	New Guinea	Bouro	14	1			chap.5 et 6
195	HAES	Papua New Guinea	Southern Highland	New Guinea	Kana	14	1			chap.5 et 6
248	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	14	1			chap.5 et 6
249	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	14	1			chap.5 et 6
24	HAES	Papua New Guinea	Southern Highland	New Guinea	Tai Ina	1_2	2			chap.5 et 6
251	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	1_2	2			chap.5 et 6
253	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	1_2	2			chap.5 et 6
254	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	14	1			chap.5 et 6
258	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	1_2	2			chap.5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
262	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	14	1			chap.5 et 6
263	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	14	1			chap.5 et 6
26	HAES	Papua New Guinea	Southern Highland	New Guinea	Pepo Ina	1_2	2			chap.5 et 6
273	HAES	Papua New Guinea	Southern Highland	New Guinea	KACC	1_2	2			chap.5 et 6
316	HAES	Papua New Guinea	Southern Highland	New Guinea	Markham	1_2	2			chap.5 et 6
318	HAES	Papua New Guinea	Southern Highland	New Guinea		14	1			chap.5 et 6
328	HAES	Papua New Guinea	Southern Highland	New Guinea	Porenong	14	1			chap.5 et 6
329	HAES	Papua New Guinea	Southern Highland	New Guinea		1_2	2			chap.5 et 6
330	HAES	Papua New Guinea	Southern Highland	New Guinea		1_2	2			chap.5 et 6
333	HAES	Papua New Guinea	Southern Highland	New Guinea		1_2	2			chap.5 et 6
338	HAES	Papua New Guinea	Southern Highland	New Guinea	Wanmun	14	1			chap.5 et 6
341	HAES	Papua New Guinea	Southern Highland	New Guinea		14	1			chap.5 et 6
343	HAES	Papua New Guinea	Southern Highland	New Guinea		1_2	2			chap.5 et 6
345	HAES	Papua New Guinea	Southern Highland	New Guinea		1_2	2			chap.5 et 6
346	HAES	Papua New Guinea	Southern Highland	New Guinea	Tul	1_2	2			chap.5 et 6
347	HAES	Papua New Guinea	Southern Highland	New Guinea	Kariap	1_2	2			chap.5 et 6
363	HAES	Papua New Guinea	Southern Highland	New Guinea	Seka Kake	1_2	2			chap.5 et 6
435	HAES	Papua New Guinea	Southern Highland	New Guinea	Peripan	14	1			chap.5 et 6
472	HAES	Papua New Guinea	Southern Highland	New Guinea	Habare	14	1			chap.5 et 6
489	HAES	Papua New Guinea	Southern Highland	New Guinea	Goroka	1_2	2			chap.5 et 6
494	HAES	Papua New Guinea	Southern Highland	New Guinea	Tari Ina	1_2	2			chap.5 et 6
500	HAES	Papua New Guinea	Southern Highland	New Guinea	Kuli Ina	1_2	2			chap.5 et 6
505	HAES	Papua New Guinea	Southern Highland	New Guinea		1_2	2			chap.5 et 6
588	HAES	Papua New Guinea	Southern Highland	New Guinea	Terek	1_2	2			chap.5 et 6
597	HAES	Papua New Guinea	Southern Highland	New Guinea	Cayman	14	1			chap.5 et 6
681	HAES	Papua New Guinea	Southern Highland	New Guinea	Koneme	1_2	2			chap.5 et 6
716	HAES	Papua New Guinea	Southern Highland	New Guinea		14	1			chap.5 et 6
723	HAES	Papua New Guinea	Southern Highland	New Guinea	Pigaig	1_2	2			chap.5 et 6
727	HAES	Papua New Guinea	Southern Highland	New Guinea	Miminbi	1_2	2			chap.5 et 6
736	HAES	Papua New Guinea	Southern Highland	New Guinea	Bone	13	1			chap.5 et 6
739	HAES	Papua New Guinea	Southern Highland	New Guinea	Sumenda	1_2	2		101	chap.5 et 6
742	HAES	Papua New Guinea	Southern Highland	New Guinea	Mundu Jame	1_2	2			chap.5 et 6
757	HAES	Papua New Guinea	Southern Highland	New Guinea	Sekamonde	1_2	2			chap.5 et 6
761	HAES	Papua New Guinea	Southern Highland	New Guinea	Rakipa	1_2	2			chap.5 et 6
769	HAES	Papua New Guinea	Southern Highland	New Guinea	Riya	1_2	2			chap.5 et 6
887	HAES	Papua New Guinea	Southern Highland	New Guinea	Annmadam	6_9	2			chap.5 et 6
339	HAES	Papua New Guinea	Southern Highland	New Guinea		14	1			chap.5 et 6
340	HAES	Papua New Guinea	Southern Highland	New Guinea		14	1			chap.5 et 6
3	LAES	Papua New Guinea	Southern Highland	New Guinea	SIL 2	1_2	2	18		chap.5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
108	LAES	Papua New Guinea	Southern Highland	New Guinea	PNG 506	1_2	2			chap.5 et 6
11	LAES	Papua New Guinea	Southern Highland	New Guinea	SIL 2	1_2	2			chap.5 et 6
111	LAES	Papua New Guinea	Southern Highland	New Guinea	ULA 2	14	1			chap.5 et 6
150	LAES	Papua New Guinea	Southern Highland	New Guinea	SIL 3	1_2	2			chap.5 et 6
224	LAES	Papua New Guinea	Southern Highland	New Guinea	LP 04 Poso village	1_2	2			chap.5 et 6
99	LAES	Papua New Guinea	Southern Highland	New Guinea	ULA 1	14	1			chap.5 et 6
970	HAES	Papua New Guinea	Western	New Guinea	Manda.A	14	1			chap.5 et 6
76	LAES	Papua New Guinea	Western	New Guinea	L 50	1_2	2			chap.5 et 6
61	HAES	Papua New Guinea	Western highland	New Guinea	Rome Amb	1_2	2	13	13	chap.5 et 6
178	HAES	Papua New Guinea	Western highland	New Guinea	Mine Donka	1_2	2	29		chap.5 et 6
419	HAES	Papua New Guinea	Western highland	New Guinea	Modamb	12	1	38	38	chap.5 et 6
441	HAES	Papua New Guinea	Western highland	New Guinea	Apuril	1_2	2			chap.5 et 6
842	HAES	Papua New Guinea	Western highland	New Guinea	Yambane	14	1			chap.5 et 6
108	HAES	Papua New Guinea	Western highland	New Guinea	Kasil	14	1			chap.5 et 6
55	HAES	Papua New Guinea	Western highland	New Guinea		14	1			chap.5 et 6
60	HAES	Papua New Guinea	Western highland	New Guinea	Jimbon Amb	1_2	2			chap.5 et 6
289	Yen	Papua New Guinea	Western Highlands, Guntz	New Guinea		14	1			chap.5 et 6
298	Yen	Papua New Guinea	Western Highlands, Kaugel V	New Guinea	mom	1_2	2	32		chap.5 et 6
304	Yen	Papua New Guinea	Western Highlands, Kaugel V	New Guinea	unju ka	1_2	2	44	44	chap.5 et 6
293	Yen	Papua New Guinea	Western Highlands, Kaugel V	New Guinea		1_2	2			chap.5 et 6
295	Yen	Papua New Guinea	Western Highlands, Kaugel V	New Guinea		1_2	2			chap.5 et 6
297	Yen	Papua New Guinea	Western Highlands, Kaugel V	New Guinea		1_2	2			chap.5 et 6
301	Yen	Papua New Guinea	Western Highlands, Kaugel V	New Guinea	tangambityə	1_2	2			chap.5 et 6
302	Yen	Papua New Guinea	Western Highlands, Kaugel V	New Guinea	arakoyə	1_2	2			chap.5 et 6
316	Yen	Papua New Guinea	Western Highlands, Kaugel V	New Guinea	mango pora	14	1			chap.5 et 6
322	Yen	Papua New Guinea	Western Highlands, Kaugel V	New Guinea	keratl	1_2	2			chap.5 et 6
327	Yen	Papua New Guinea	Western Highlands, Kaugel V	New Guinea	simbu angu	14	1			chap.5 et 6
337	Yen	Papua New Guinea	Western Highlands, Kaugel V	New Guinea	konmay	14	1			chap.5 et 6
338	Yen	Papua New Guinea	Western Highlands, Kaugel V	New Guinea	bajj	5	1			chap.5 et 6
341	Yen	Papua New Guinea	Western Highlands, Kaugel V	New Guinea	acic	14	1			chap.5 et 6
210	Yen	West New Guinea	Wissellakes	New Guinea	notta dere	1_2	2	33		chap.5 et 6
211	Yen	West New Guinea	Wissellakes	New Guinea	notta molodoer	1_2	2	33		chap.5 et 6
212	Yen	West New Guinea	Wissellakes	New Guinea	notta waikio	ND	ND			chap.5 et 6
440131	CIP	Papua New Guinea		New Guinea		14	1			chap.5 et 6
440309	CIP	Papua New Guinea		New Guinea		13	1			chap.5 et 6
440422	CIP	Papua New Guinea		New Guinea		14	1			chap.5 et 6
440808	CIP	Papua New Guinea		New Guinea		1_2	2			chap.5 et 6
440222	CIP	West New Guinea		New Guinea		1_2	2			chap.5 et 6
440283	CIP	West New Guinea		New Guinea		14	1			chap.5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
440408	CIP	West New Guinea		New Guinea		1_2	2			chap.5 et 6
440459	CIP	West New Guinea		New Guinea		14	1			chap.5 et 6
440479	CIP	West New Guinea		New Guinea		14	1			chap.5 et 6
440480	CIP	West New Guinea		New Guinea		1_2	2			chap.5 et 6
440481	CIP	West New Guinea		New Guinea		13	1			chap.5 et 6
440482	CIP	West New Guinea		New Guinea		1_2	2			chap.5 et 6
443641	CIP	West New Guinea		New Guinea		14	1			chap.5 et 6
443647	CIP	West New Guinea		New Guinea		14	1			chap.5 et 6
443650	CIP	West New Guinea		New Guinea		1_2	2			chap.5 et 6
443653	CIP	West New Guinea		New Guinea		1_2	2			chap.5 et 6
443654	CIP	West New Guinea		New Guinea		14	1			chap.5 et 6
443658	CIP	West New Guinea		New Guinea		14	1			chap.5 et 6
443663	CIP	West New Guinea		New Guinea		1_2	2			chap.5 et 6
443677	CIP	West New Guinea		New Guinea		6_9	2			chap.5 et 6
443681	CIP	West New Guinea		New Guinea		6_9	2			chap.5 et 6
443689	CIP	West New Guinea		New Guinea		1_2	2			chap.5 et 6
443695	CIP	West New Guinea		New Guinea		1_2	2			chap.5 et 6
443697	CIP	West New Guinea		New Guinea		6_9	2			chap.5 et 6
443700	CIP	West New Guinea		New Guinea		1_2	2			chap.5 et 6
443708	CIP	West New Guinea		New Guinea		1_2	2			chap.5 et 6
443728	CIP	West New Guinea		New Guinea		1_2	2			chap.5 et 6
443730	CIP	West New Guinea		New Guinea		1_2	2			chap.5 et 6
443734	CIP	West New Guinea		New Guinea		6_9	2			chap.5 et 6
443748	CIP	West New Guinea		New Guinea		14	1			chap.5 et 6
443749	CIP	West New Guinea		New Guinea		6_9	2			chap.5 et 6
12_EN	WS	West New Guinea		New Guinea		33	2			chap.5 et 6
3_EN	WS	West New Guinea		New Guinea		16_23	2			chap.5 et 6
1_EN	WS	West New Guinea		New Guinea		26	1			chap.5 et 6
10_EN	WS	West New Guinea		New Guinea		14	1			chap.5 et 6
11_EN	WS	West New Guinea		New Guinea		6_9	2			chap.5 et 6
13_EN	WS	West New Guinea		New Guinea		14	1			chap.5 et 6
14_EN	WS	West New Guinea		New Guinea		14	1			chap.5 et 6
15_EN	WS	West New Guinea		New Guinea		25	2			chap.5 et 6
2_EN	WS	West New Guinea		New Guinea		14	1			chap.5 et 6
5_EN	WS	West New Guinea		New Guinea		1_2	2			chap.5 et 6
7_EN	WS	West New Guinea		New Guinea		26	1			chap.5 et 6
9_EN	WS	West New Guinea		New Guinea		14	1			chap.5 et 6
4_EN	WS	West New Guinea		New Guinea		1_2	2			chap.5 et 6
6_EN	WS	West New Guinea		New Guinea		6_9	2			chap.5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
401114	CIP	Guatemala	Alta Verapaz	Northern region	Is	14	1			chap.3, 5 et 6
401077	CIP	Guatemala	Alta Verapaz	Northern region	Camote Morado	6_9	2			chap.3, 5 et 6
401129	CIP	Guatemala	Alta Verapaz	Northern region	Camote Blanco	6_9	2			chap.3, 5 et 6
401172	CIP	Jamaica	An	Northern region	Browns Town	1_2	2	60		chap.3, 5 et 6
401170	CIP	Jamaica	An	Northern region	Watson	1_2	2			chap.3, 5 et 6
401169	CIP	Jamaica	An	Northern region	Herbie	1_2	2			chap.3, 5 et 6
400377	CIP	Columbia	Atlantico	Northern region		16_23	2			chap.3, 5 et 6
401150	CIP	Honduras	Atlantida	Northern region	Dos Semanas	6_9	2			chap.3, 5 et 6
401151	CIP	Honduras	Atlantida	Northern region	Morado	1_2	2			chap.3, 5 et 6
401152	CIP	Honduras	Atlantida	Northern region	Rojo	6_9	2			chap.3, 5 et 6
401153	CIP	Honduras	Atlantida	Northern region	Amarillo	1_2	2			chap.3, 5 et 6
401119	CIP	Guatemala	Baja Verapaz	Northern region	Camote Morado	13	1			chap.3, 5 et 6
401122	CIP	Guatemala	Baja Verapaz	Northern region	Camote Blanco	6_9	2			chap.3, 5 et 6
400376	CIP	Columbia	Bolivar	Northern region	Amarilla	16_23	2			chap.3, 5 et 6
400379	CIP	Columbia	Bolivar	Northern region	Morada Blanca	16_23	2			chap.3, 5 et 6
401200	CIP	Jamaica	Catherin	Northern region	Rhoden	14	1			chap.3, 5 et 6
401194	CIP	Jamaica	Catherin	Northern region	Six Weeks	1_2	2			chap.3, 5 et 6
401197	CIP	Jamaica	Catherin	Northern region	Hawk	1_2	2			chap.3, 5 et 6
441777	CIP	Jamaica	Catherin	Northern region	Springfield	1_2	2			chap.3, 5 et 6
441778	CIP	Jamaica	Catherin	Northern region	LA 35	3	2			chap.3, 5 et 6
401551	CIP	Columbia	Cauca	Northern region		4	1	54		chap.3, 5 et 6
400509	CIP	Columbia	Cesar	Northern region	blanca	14	1	54		chap.3, 5 et 6
400518	CIP	Columbia	Cesar	Northern region	batata Amarilla	14	1	54		chap.3, 5 et 6
400512	CIP	Columbia	Cesar	Northern region		24	2			chap.3, 5 et 6
400504	CIP	Columbia	Cesar	Northern region	amarilla	24	2			chap.3, 5 et 6
400508	CIP	Columbia	Cesar	Northern region		6_9	2			chap.3, 5 et 6
400511	CIP	Columbia	Cesar	Northern region	Blanca	1_2	2			chap.3, 5 et 6
400822	CIP	Dominican Republic	cey Al Medio	Northern region	Canabacoa	1_2	2			chap.3, 5 et 6
401205	CIP	Mexico	Chiapas	Northern region		1_2	2			chap.3, 5 et 6
401105	CIP	Guatemala	Chimaltenango	Northern region	Camote Morado	6_9	2			chap.3, 5 et 6
401107	CIP	Guatemala	Chimaltenango	Northern region		6_9	2			chap.3, 5 et 6
401255	CIP	Panama	Chiriqui	Northern region	Batatilla	1_2	2			chap.3, 5 et 6
401257	CIP	Panama	Chiriqui	Northern region	Camote	16_23	2			chap.3, 5 et 6
401183	CIP	Jamaica	Clarendon	Northern region	Horse Money	1_2	2	60		chap.3, 5 et 6
400040	CIP	Jamaica	Clarendon	Northern region	Manchester Hawk	1_2	2			chap.3, 5 et 6
401184	CIP	Jamaica	Clarendon	Northern region	Sigies Short Internode	1_2	2			chap.3, 5 et 6
401190	CIP	Jamaica	Clarendon	Northern region	Pearl	3	2			chap.3, 5 et 6
401261	CIP	Panama	Cocle	Northern region	Morada	16_23	2			chap.3, 5 et 6
401378	CIP	Venezuela	Cojedes	Northern region	Rosada	1_2	2			chap.3, 5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
400356	CIP	Columbia	Cordoba	Northern region	Amarillo	16_23	2			chap.3, 5 et 6
400360	CIP	Columbia	Cordoba	Northern region		16_23	2			chap.3, 5 et 6
400364	CIP	Columbia	Cordoba	Northern region	de Leche	7	2			chap.3, 5 et 6
400367	CIP	Columbia	Cordoba	Northern region		7	2			chap.3, 5 et 6
400370	CIP	Columbia	Cordoba	Northern region	Roja	16_23	2			chap.3, 5 et 6
400372	CIP	Columbia	Cordoba	Northern region	Morado	16_23	2			chap.3, 5 et 6
401154	CIP	Honduras	Cortes	Northern region		1_2	2			chap.3, 5 et 6
401267	CIP	Panama	Darien	Northern region	Blanco	16_23	2	62		chap.3, 5 et 6
401269	CIP	Panama	Darien	Northern region		16_23	2	62		chap.3, 5 et 6
401265	CIP	Panama	Darien	Northern region	Morada	22	2			chap.3, 5 et 6
401266	CIP	Panama	Darien	Northern region		16_23	2			chap.3, 5 et 6
401272	CIP	Panama	Darien	Northern region	Colorado	22	2			chap.3, 5 et 6
671	Yen	Columbia	Don Matias	Northern region	paramonguino	6_9	2			chap.3, 5 et 6
401108	CIP	Guatemala	El Progreso	Northern region	Camote Morado	1_2	2			chap.3, 5 et 6
672	Yen	Columbia	Felidia	Northern region	paramonguino	6_9	2			chap.3, 5 et 6
401431	CIP	Venezuela	Guarico	Northern region	Batata Morada	6_9	2			chap.3, 5 et 6
401104	CIP	Guatemala	Guatemala	Northern region	Camote Naranja	1_2	2			chap.3, 5 et 6
401238	CIP	Panama	Herrera	Northern region	Batata	16_23	2			chap.3, 5 et 6
401546	CIP	Columbia	Huila	Northern region		6_9	2			chap.3, 5 et 6
401101	CIP	Guatemala	Izabal	Northern region	Camote Rojo	1_2	2	57		chap.3, 5 et 6
401071	CIP	Guatemala	Izabal	Northern region	Camote Morado	6_9	2			chap.3, 5 et 6
401076	CIP	Guatemala	Izabal	Northern region	Camote Amarillo	1_2	2			chap.3, 5 et 6
401100	CIP	Guatemala	Izabal	Northern region	Camote Blanco	1_2	2			chap.3, 5 et 6
401582	CIP	Guatemala	Izabal	Northern region		1_2	2			chap.3, 5 et 6
401055	CIP	Guatemala	Jutiapa	Northern region	Camote Blanco	13	1			chap.3, 5 et 6
401056	CIP	Guatemala	Jutiapa	Northern region	Camote Blanco	1_2	2			chap.3, 5 et 6
400496	CIP	Columbia	La Guajira	Northern region		24	2			chap.3, 5 et 6
401381	CIP	Venezuela	Lara	Northern region		6_9	2			chap.3, 5 et 6
401383	CIP	Venezuela	Lara	Northern region		6_9	2			chap.3, 5 et 6
401243	CIP	Panama	Los Santos	Northern region	Amarilla	1_2	2			chap.3, 5 et 6
702	Yen	United States	Louisiana	Northern region	Myers Early	14	1	73		chap.3, 5 et 6
705	Yen	United States	Louisiana	Northern region	Acadian	14	1	74		chap.3, 5 et 6
704	Yen	United States	Louisiana	Northern region	seedling 12	14	1	74		chap.3, 5 et 6
400381	CIP	Columbia	Magdalena	Northern region	rosada	24	2			chap.3, 5 et 6
400493	CIP	Columbia	Magdalena	Northern region	batatas	7	2			chap.3, 5 et 6
401228	CIP	Nicaragua	Managua	Northern region	Batata Morada	1_2	2			chap.3, 5 et 6
401173	CIP	Jamaica	Maschester	Northern region	Amos	1_2	2			chap.3, 5 et 6
401403	CIP	Venezuela	Merida	Northern region	Morado	6_9	2	63		chap.3, 5 et 6
401501	CIP	Venezuela	Merida	Northern region		6_9	2	63		chap.3, 5 et 6



Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
401500	CIP	Venezuela	Merida	Northern region		4	1			chap.3, 5 et 6
401396	CIP	Venezuela	Merida	Northern region		6_9	2			chap.3, 5 et 6
401399	CIP	Venezuela	Merida	Northern region		14	1			chap.3, 5 et 6
401430	CIP	Venezuela	Miranda	Northern region	Morada	6_9	2			chap.3, 5 et 6
401428	CIP	Venezuela	Miranda	Northern region		16_23	2			chap.3, 5 et 6
401429	CIP	Venezuela	Miranda	Northern region	Morada	1_2	2			chap.3, 5 et 6
401445	CIP	Venezuela	Monagas	Northern region	Blanca	14	1	64		chap.3, 5 et 6
401451	CIP	Venezuela	Monagas	Northern region	Blanco	14	1	64		chap.3, 5 et 6
401454	CIP	Venezuela	Monagas	Northern region	Blanco	14	1			chap.3, 5 et 6
401455	CIP	Venezuela	Monagas	Northern region	Blanco	16_23	2			chap.3, 5 et 6
401453	CIP	Venezuela	Monagas	Northern region	Morado	1_2	2			chap.3, 5 et 6
401444	CIP	Venezuela	Monagas	Northern region	Chaco Rosado	1_2	2			chap.3, 5 et 6
401446	CIP	Venezuela	Monagas	Northern region	Blanca	16_23	2			chap.3, 5 et 6
400515	CIP	Columbia	Norte Santander	Northern region	Tuna	16_23	2			chap.3, 5 et 6
401559	CIP	Columbia	Norte Santander	Northern region		10	2			chap.3, 5 et 6
400517	CIP	Columbia	Norte Santander	Northern region	de La Fina	14	1			chap.3, 5 et 6
400521	CIP	Columbia	Norte Santander	Northern region	Amarilla Ubre De Vaca	16_23	2			chap.3, 5 et 6
400524	CIP	Columbia	Norte Santander	Northern region		16_23	2			chap.3, 5 et 6
400828	CIP	Dominican Republic	nsenor Nouel	Northern region	Andariega B	14	1	55		chap.3, 5 et 6
400830	CIP	Dominican Republic	nsenor Nouel	Northern region	Hoja de Pancho	14	1	55		chap.3, 5 et 6
401543	CIP	Dominican Republic	nsenor Nouel	Northern region	Mocana Cino	6_9	2			chap.3, 5 et 6
401560	CIP	Venezuela	Nueva Esparta	Northern region		14	1			chap.3, 5 et 6
679	Yen	Columbia	Palmira	Northern region	paramonguino	6_9	2	72		chap.3, 5 et 6
677	Yen	Columbia	Palmira	Northern region	paramonguino	ND	ND	93	93	chap.3, 5 et 6
682_1	Yen	Columbia	Palmira	Northern region	paramonguino	ND	ND			chap.3, 5 et 6
401277	CIP	Panama	Panama	Northern region	Rojokinit	14	1	50	50	chap.3, 5 et 6
401280	CIP	Panama	Panama	Northern region		14	1			chap.3, 5 et 6
401264	CIP	Panama	Panama	Northern region	Batata Roja	22	2			chap.3, 5 et 6
401236	CIP	Panama	Panama	Northern region	Blanco	14	1			chap.3, 5 et 6
401263	CIP	Panama	Panama	Northern region	Morado	22	2			chap.3, 5 et 6
401094	CIP	Guatemala	Peten	Northern region	Camote Amarillo	17	2			chap.3, 5 et 6
401062	CIP	Guatemala	Peten	Northern region	Camote Moral	6_9	2			chap.3, 5 et 6
401067	CIP	Guatemala	Peten	Northern region	Camote Morado	1_2	2			chap.3, 5 et 6
401087	CIP	Guatemala	Peten	Northern region	Camote Morado	6_9	2			chap.3, 5 et 6
401088	CIP	Guatemala	Peten	Northern region	Camote Blanco	6_9	2			chap.3, 5 et 6
401089	CIP	Guatemala	Peten	Northern region	Camote Injerto	17	2			chap.3, 5 et 6
401092	CIP	Guatemala	Peten	Northern region	Camote Blanco	1_2	2			chap.3, 5 et 6
401095	CIP	Guatemala	Peten	Northern region	Camote Morado	1_2	2			chap.3, 5 et 6
401053	CIP	Guatemala	Progreso	Northern region	Camote Morado	1_2	2			chap.3, 5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
401587	CIP	Guatemala	Quiche	Northern region		14	1	50		chap.3, 5 et 6
401541	CIP	Dominican Republic	ravia	Northern region	Colorado	6_9	2			chap.3, 5 et 6
400817	CIP	Dominican Republic	ravia	Northern region	Varga	14	1			chap.3, 5 et 6
400819	CIP	Dominican Republic	ravia	Northern region	Tunita	6_9	2			chap.3, 5 et 6
686	Yen	Columbia	Retiro de de los Indios	Northern region	paramonguino	16_23	2			chap.3, 5 et 6
688	Yen	Columbia	Retiro de de los Indios	Northern region	paramonguino	16_23	2			chap.3, 5 et 6
687	Yen	Columbia	Retiro de de los Indios	Northern region	paramonguino	16_23	2			chap.3, 5 et 6
401224	CIP	Nicaragua	Rio San Juan	Northern region	Camote	1_2	2			chap.3, 5 et 6
401225	CIP	Nicaragua	Rio San Juan	Northern region	Camote	6_9	2			chap.3, 5 et 6
401079	CIP	Guatemala	Sacatepequez	Northern region	Camote Blanco	6_9	2			chap.3, 5 et 6
401111	CIP	Guatemala	San Marcos	Northern region	Camote Morado	1_2	2			chap.3, 5 et 6
401083	CIP	Guatemala	Santa Rosa	Northern region	Camote	1_2	2			chap.3, 5 et 6
401136	CIP	Guatemala	Solola	Northern region	Is Blanco	1_2	2			chap.3, 5 et 6
400375	CIP	Columbia	Sucre	Northern region	Mariquitena	16_23	2			chap.3, 5 et 6
401436	CIP	Venezuela	Sucre	Northern region	Blanco	1_2	2			chap.3, 5 et 6
401437	CIP	Venezuela	Sucre	Northern region	Chaco Blanco	14	1			chap.3, 5 et 6
401432	CIP	Venezuela	Sucre	Northern region	Batata Rosa	1_2	2			chap.3, 5 et 6
401433	CIP	Venezuela	Sucre	Northern region	Cerro Colorado	16_23	2			chap.3, 5 et 6
401434	CIP	Venezuela	Sucre	Northern region	Chaco Morado	6_9	2			chap.3, 5 et 6
401435	CIP	Venezuela	Sucre	Northern region		1_2	2			chap.3, 5 et 6
401439	CIP	Venezuela	Sucre	Northern region	Chaco Rosado	16_23	2			chap.3, 5 et 6
401442	CIP	Venezuela	Sucre	Northern region	Batata Morada	1_2	2			chap.3, 5 et 6
401464	CIP	Venezuela	Trujillo	Northern region	Patatilla	6_9	2	63		chap.3, 5 et 6
401394	CIP	Venezuela	Trujillo	Northern region	Barbechera	6_9	2			chap.3, 5 et 6
401392	CIP	Venezuela	Trujillo	Northern region	Carachera	6_9	2			chap.3, 5 et 6
400825	CIP	Dominican Republic	Vega	Northern region	La Cheche Moradita	6_9	2			chap.3, 5 et 6
401251	CIP	Panama	Veraguas	Northern region	Morado	16_23	2	62		chap.3, 5 et 6
401259	CIP	Panama	Veraguas	Northern region	Roja	16_23	2	62		chap.3, 5 et 6
401248	CIP	Panama	Veraguas	Northern region	Amarillo	1_2	2			chap.3, 5 et 6
401246	CIP	Panama	Veraguas	Northern region		16_23	2			chap.3, 5 et 6
401249	CIP	Panama	Veraguas	Northern region	Morada	1_2	2			chap.3, 5 et 6
401252	CIP	Panama	Veraguas	Northern region		1_2	2			chap.3, 5 et 6
401253	CIP	Panama	Veraguas	Northern region	Camote	1_2	2			chap.3, 5 et 6
400551	CIP	Costa Rica		Northern region		14	1			chap.3, 5 et 6
400543	CIP	Costa Rica		Northern region	Variedad Local	1_2	2			chap.3, 5 et 6
400608	CIP	Cuba		Northern region	Haiti	14	1			chap.3, 5 et 6
400578	CIP	Cuba		Northern region	Aparecido Manchado	6_9	2			chap.3, 5 et 6
400584	CIP	Cuba		Northern region	Bonito	14	1			chap.3, 5 et 6
400602	CIP	Cuba		Northern region	Engorda Villareno	6_9	2			chap.3, 5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
400589	CIP	Cuba		Northern region	Camarioca	3	2			chap.3, 5 et 6
400615	CIP	Cuba		Northern region	Martinica de Matojo	3	2			chap.3, 5 et 6
400621	CIP	Cuba		Northern region	Palenque	6_9	2			chap.3, 5 et 6
400622	CIP	Cuba		Northern region	Pan Con Vino	1_2	2			chap.3, 5 et 6
400628	CIP	Cuba		Northern region	Peralejo Rojo	6_9	2			chap.3, 5 et 6
400632	CIP	Cuba		Northern region	Santiaguero	3	2			chap.3, 5 et 6
400034	CIP	Dominican Republic		Northern region		1_2	2	53		chap.3, 5 et 6
400865	CIP	Dominican Republic		Northern region	Celestina	1_2	2	53		chap.3, 5 et 6
400786	CIP	Dominican Republic		Northern region	Copela	14	1			chap.3, 5 et 6
400029	CIP	Dominican Republic		Northern region	Tela de Cebolla	1_2	2			chap.3, 5 et 6
400796	CIP	Dominican Republic		Northern region	Sola No Estan	1_2	2			chap.3, 5 et 6
400808	CIP	Dominican Republic		Northern region		1_2	2			chap.3, 5 et 6
400838	CIP	Dominican Republic		Northern region	Roja de La Reforma	1_2	2			chap.3, 5 et 6
400846	CIP	Dominican Republic		Northern region	Envicia Raton	1_2	2			chap.3, 5 et 6
400847	CIP	Dominican Republic		Northern region	Pata De Gallina	1_2	2			chap.3, 5 et 6
400849	CIP	Dominican Republic		Northern region	Cano	1_2	2			chap.3, 5 et 6
400852	CIP	Dominican Republic		Northern region	Mata Hambre	1_2	2			chap.3, 5 et 6
400853	CIP	Dominican Republic		Northern region	Yema de Huevo	1_2	2			chap.3, 5 et 6
400866	CIP	Dominican Republic		Northern region	La Blanca	3	2			chap.3, 5 et 6
401214	CIP	Mexico		Northern region	Criollo Cacahuatan,chis.	1_2	2	61		chap.3, 5 et 6
401215	CIP	Mexico		Northern region	Coleccion Tierra Blanca	1_2	2	61		chap.3, 5 et 6
401475	CIP	Mexico		Northern region		14	1	104	104	chap.3, 5 et 6
401207	CIP	Mexico		Northern region	Santa Sofia	6_9	2			chap.3, 5 et 6
441724	CIP	Mexico		Northern region	Cuitzeo	1_2	2			chap.3, 5 et 6
401208	CIP	Mexico		Northern region	Mesquita	6_9	2			chap.3, 5 et 6
400009	CIP	Mexico		Northern region		14	1			chap.3, 5 et 6
400010	CIP	Mexico		Northern region	226	1_2	2			chap.3, 5 et 6
401213	CIP	Mexico		Northern region	Tinto	6_9	2			chap.3, 5 et 6
401211	CIP	Mexico		Northern region	Criollo de Veracruz	1_2	2			chap.3, 5 et 6
401212	CIP	Mexico		Northern region	Regional de Tehuantepec	1_2	2			chap.3, 5 et 6
401216	CIP	Mexico		Northern region	Coleccion Bajo Papaloapar	6_9	2			chap.3, 5 et 6
401217	CIP	Mexico		Northern region	PI 267946	1_2	2			chap.3, 5 et 6
401222	CIP	Mexico		Northern region	Cuitzo Col. 8463	1_2	2			chap.3, 5 et 6
401468	CIP	Mexico		Northern region		16_23	2			chap.3, 5 et 6
401470	CIP	Mexico		Northern region		18	2			chap.3, 5 et 6
401471	CIP	Mexico		Northern region		6_9	2			chap.3, 5 et 6
401472	CIP	Mexico		Northern region		6_9	2			chap.3, 5 et 6
404018	CIP	Mexico		Northern region		3	2			chap.3, 5 et 6
404019	CIP	Mexico		Northern region		3	2			chap.3, 5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
404021	CIP	Mexico		Northern region		3	2			chap.3, 5 et 6
404024	CIP	Mexico		Northern region		1_2	2			chap.3, 5 et 6
404026	CIP	Mexico		Northern region		6_9	2			chap.3, 5 et 6
404027	CIP	Mexico		Northern region		1_2	2			chap.3, 5 et 6
440052	CIP	Puerto Rico		Northern region	Margarita (SPV 70)	1_2	2			chap.3, 5 et 6
440206	CIP	Puerto Rico		Northern region		1_2	2			chap.3, 5 et 6
400025	CIP	Saint Vincent		Northern region	Lovers Name	14	1	104	104	chap.3, 5 et 6
400024	CIP	Saint Vincent		Northern region	Barbados	1_2	2			chap.3, 5 et 6
400026	CIP	Saint Vincent		Northern region	Rhode Vine	1_2	2			chap.3, 5 et 6
710	Yen	Mexico		Northern region		1_2	2			chap.3, 5 et 6
712	Yen	Puerto Rico		Northern region		ND	ND	73		chap.3, 5 et 6
482	Yen	Cook Islands	Aitutaki	Polynesia	tauranga/torumarama	1_2	2			chap.6
502	Yen	New Zealand	Auckland	Polynesia	gisborne red	14	1	2	2	chap.6
501_1	Yen	New Zealand	Auckland	Polynesia	tauranga red	14	1	67	67	chap.6
501_2	Yen	New Zealand	Auckland	Polynesia	tauranga red	14	1			chap.6
500	Yen	New Zealand	Auckland	Polynesia	owairaka red	14	1			chap.6
533	Yen	Society Islands	Borabora	Polynesia	umara hererei	1_2	2	1	1B	chap.6
532	Yen	Society Islands	Borabora	Polynesia	umara vario	1_2	2			chap.6
553	Yen	Marquesas Islands	Fatuhiva	Polynesia	a'poa	1_2	2	1		chap.6
570_2	Yen	Easter Island	Hangaroa	Polynesia	arenga riki ri	14	1			chap.6
Ta_33	SDR	Marquesas Islands	Hiva Oa	Polynesia		1_2	2	1		chap.6
Ta_35	SDR	Marquesas Islands	Hiva Oa	Polynesia		1_2	2	1		chap.6
Ta_28	SDR	Marquesas Islands	Hiva Oa	Polynesia		6_9	2	78	78B	chap.6
Ta_23	SDR	Marquesas Islands	Hiva Oa	Polynesia		1_2	2			chap.6
Ta_34	SDR	Marquesas Islands	Hiva Oa	Polynesia		14	1			chap.6
Ta_29	SDR	Marquesas Islands	Hiva Oa	Polynesia		6_9	2			chap.6
Ta_15	SDR	Marquesas Islands	Hiva Oa	Polynesia		1_2	2			chap.6
Ta_30	SDR	Marquesas Islands	Hiva Oa	Polynesia		30	1			chap.6
Ta_31	SDR	Marquesas Islands	Hiva Oa	Polynesia		1_2	2			chap.6
Ta_32	SDR	Marquesas Islands	Hiva Oa	Polynesia		1_2	2			chap.6
Ta_64	SDR	Society Islands	Huahine	Polynesia		30	1	106		chap.6
Ta_63	SDR	Society Islands	Huahine	Polynesia		30	1			chap.6
Ta_62	SDR	Society Islands	Huahine	Polynesia		1_2	2			chap.6
Ta_65	SDR	Society Islands	Huahine	Polynesia		30	1			chap.6
507	Yen	New Zealand	Kawhia	Polynesia	kotepo	14	1	2		chap.6
491	Yen	Cook Islands	Mangaia	Polynesia		1_2	2	1		chap.6
490	Yen	Cook Islands	Mangaia	Polynesia		14	1	67	67	chap.6
525	Yen	Society Islands	Moorea	Polynesia		1_2	2	1	1B	chap.6
544	Yen	Society Islands	Moorea	Polynesia		1_2	2	1	1B	chap.6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
528	Yen	Society Islands	Moorea	Polynesia		1_2	2	1		chap.6
545	Yen	Society Islands	Moorea	Polynesia	umara re'amoa	1_2	2			chap.6
Ta_26	SDR	Marquesas Islands	Nuku Hiva	Polynesia	Eka mamaimoa	1_2	2			chap.6
Ta_27	SDR	Marquesas Islands	Nuku Hiva	Polynesia	Arei	1_2	2			chap.6
Ta_24	SDR	Marquesas Islands	Nuku Hiva	Polynesia	tokatoka	6_9	2			chap.6
Ta_25	SDR	Marquesas Islands	Nuku Hiva	Polynesia	Roi	1_2	2			chap.6
443	Yen	Tonga	Nukua' lofa	Polynesia	kumala lau manioke	1_2	2	66		chap.6
452A	Yen	Tonga	Nukua' lofa	Polynesia	utation of Y452/ Tolumahi	31	1			chap.6
452	Yen	Tonga	Nukua' lofa	Polynesia	vava	14	1			chap.6
458	Yen	Tonga	Nukua' lofa	Polynesia		1_2	2			chap.6
452B	Yen	Tonga	Nukua' lofa	Polynesia	utation of Y452/ Tolumahi	14	1			chap.6
556	Yen	Marquesas Islands	Nukuhiva	Polynesia	papau	1_2	2	1		chap.6
557	Yen	Marquesas Islands	Nukuhiva	Polynesia	pikino	14	1			chap.6
Ta_37	SDR	Society Islands	Raiatea	Polynesia		6_9	2	78	78B	chap.6
Ta_43	SDR	Society Islands	Raiatea	Polynesia		6_9	2	78	78B	chap.6
Ta_40	SDR	Society Islands	Raiatea	Polynesia		1_2	2	85	85	chap.6
Ta_56	SDR	Society Islands	Raiatea	Polynesia		1_2	2	85	85	chap.6
Ta_53	SDR	Society Islands	Raiatea	Polynesia		1_2	2	87		chap.6
Ta_47	SDR	Society Islands	Raiatea	Polynesia		1_2	2	87		chap.6
Ta_52	SDR	Society Islands	Raiatea	Polynesia		1_2	2	88		chap.6
Ta_59	SDR	Society Islands	Raiatea	Polynesia		1_2	2	88		chap.6
Ta_49	SDR	Society Islands	Raiatea	Polynesia		30	1	106		chap.6
Ta_61	SDR	Society Islands	Raiatea	Polynesia		1_2	2			chap.6
Ta_39	SDR	Society Islands	Raiatea	Polynesia		14	1			chap.6
Ta_45	SDR	Society Islands	Raiatea	Polynesia		1_2	2			chap.6
Ta_42	SDR	Society Islands	Raiatea	Polynesia		6_9	2			chap.6
Ta_38	SDR	Society Islands	Raiatea	Polynesia		1_2	2			chap.6
Ta_36	SDR	Society Islands	Raiatea	Polynesia		1_2	2			chap.6
Ta_60	SDR	Society Islands	Raiatea	Polynesia		1_2	2			chap.6
Ta_48	SDR	Society Islands	Raiatea	Polynesia		6_9	2			chap.6
Ta_55	SDR	Society Islands	Raiatea	Polynesia		1_2	2			chap.6
Ta_57	SDR	Society Islands	Raiatea	Polynesia		6_9	2			chap.6
Ta_58	SDR	Society Islands	Raiatea	Polynesia		14	1			chap.6
Ta_41	SDR	Society Islands	Raiatea	Polynesia		14	1			chap.6
Ta_44	SDR	Society Islands	Raiatea	Polynesia		6_9	2			chap.6
Ta_46	SDR	Society Islands	Raiatea	Polynesia		6_9	2			chap.6
Ta_50	SDR	Society Islands	Raiatea	Polynesia		1_2	2			chap.6
Ta_51	SDR	Society Islands	Raiatea	Polynesia		1_2	2			chap.6
Ta_54	SDR	Society Islands	Raiatea	Polynesia		1_2	2			chap.6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
537	Yen	Society Islands	Raiatea	Polynesia	umara pitanya	1_2	2	1		chap.6
539	Yen	Society Islands	Raiatea	Polynesia		1_2	2			chap.6
Ta_1	SDR	Austral Islands	Raivavae	Polynesia	Raiavavae	1_2	2	1	1B	chap.6
Ta_3	SDR	Austral Islands	Raivavae	Polynesia	Rea rea	6_9	2	78		chap.6
Ta_14	SDR	Austral Islands	Raivavae	Polynesia	Vareau	1_2	2			chap.6
Ta_5	SDR	Austral Islands	Raivavae	Polynesia	Ute ute Tubuai	1_2	2			chap.6
Ta_6	SDR	Austral Islands	Raivavae	Polynesia	Ute ute	1_2	2			chap.6
Ta_2	SDR	Austral Islands	Raivavae	Polynesia	Ute ute	30	1			chap.6
Ta_13	SDR	Austral Islands	Raivavae	Polynesia		1_2	2			chap.6
Ta_4	SDR	Austral Islands	Raivavae	Polynesia		1_2	2			chap.6
485	Yen	Cook Islands	Rarotonga	Polynesia	rokini 1	1_2	2	1		chap.6
Ta_21	SDR	Austral Islands	Rimatara	Polynesia		6_9	2	78	78B	chap.6
Ta_19	SDR	Austral Islands	Rimatara	Polynesia		1_2	2	84		chap.6
Ta_20	SDR	Austral Islands	Rimatara	Polynesia		1_2	2	84		chap.6
Ta_18	SDR	Austral Islands	Rimatara	Polynesia		1_2	2			chap.6
Ta_17	SDR	Austral Islands	Rimatara	Polynesia		30	1			chap.6
Ta_22	SDR	Austral Islands	Rimatara	Polynesia		1_2	2			chap.6
542	Yen	Society Islands	Tahiti	Polynesia		1_2	2	1	1B	chap.6
513_2	Yen	New Zealand	Taputini	Polynesia	paihia	14	1	68		chap.6
503	Yen	New Zealand	Taranaki	Polynesia	waina	14	1	2	2	chap.6
512	Yen	New Zealand	Tauranga	Polynesia	utation of Y511 (Huti Hut	14	1	67		chap.6
508	Yen	New Zealand	Tauranga	Polynesia	rekamaroa	14	1	68		chap.6
Ta_11	SDR	Marquesas Islands	Ua Pou	Polynesia	Pikihariete	1_2	2	1		chap.6
Ta_10	SDR	Marquesas Islands	Ua Pou	Polynesia	Hahaavai	6_9	2	78		chap.6
Ta_12	SDR	Marquesas Islands	Ua Pou	Polynesia	Apoa	6_9	2	78		chap.6
Ta_7	SDR	Marquesas Islands	Ua Pou	Polynesia	Huearanahi	14	1			chap.6
Ta_16	SDR	Marquesas Islands	Ua Pou	Polynesia	Hakitao	30	1			chap.6
Ta_8	SDR	Marquesas Islands	Ua Pou	Polynesia	Ria	6_9	2			chap.6
Ta_9	SDR	Marquesas Islands	Ua Pou	Polynesia	Kovake	6_9	2			chap.6
465	Yen	Samoa	Upolu	Polynesia		1_2	2			chap.6
466	Yen	Samoa	Upolu	Polynesia		1_2	2			chap.6
561	Yen	Marquesas Islands	Vapau	Polynesia	a'poa	1_2	2	70		chap.6
563	Yen	Marquesas Islands	Vapau	Polynesia	hou	1_2	2			chap.6
440294	CIP	Cook Islands		Polynesia		14	1			chap.6
440448	CIP	Cook Islands		Polynesia		14	1			chap.6
440398	CIP	New Zealand		Polynesia		14	1			chap.6
440643	CIP	Niue Island		Polynesia		14	1			chap.6
440273	CIP	Tonga		Polynesia		14	1			chap.6
440274	CIP	Tonga		Polynesia		6_9	2			chap.6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
440275	CIP	Tonga		Polynesia		14	1			chap.6
440277	CIP	Tonga		Polynesia		1_2	2			chap.6
440280	CIP	Tonga		Polynesia		1_2	2			chap.6
440281	CIP	Tonga		Polynesia		14	1			chap.6
440282	CIP	Tonga		Polynesia		1_2	2			chap.6
440373	CIP	Tonga		Polynesia		6_9	2			chap.6
440374	CIP	Tonga		Polynesia		14	1			chap.6
440384	CIP	Tonga		Polynesia		1_2	2			chap.6
HLA_1	HLA	Hawaii		Polynesia	kala	6_9	2	78	78A	chap.6
HLA_2	HLA	Hawaii		Polynesia	piko	30	1	79		chap.6
HLA_5	HLA	Hawaii		Polynesia	huawowa	14	1	83		chap.6
HLA_3	HLA	Hawaii		Polynesia	kalia	14	1	83	83	chap.6
HLA_4	HLA	Hawaii		Polynesia	pauiaaka	14	1	83	83	chap.6
HLA_8	HLA	Hawaii		Polynesia	palau	1_2	2			chap.6
HLA_9	HLA	Hawaii		Polynesia	mohihi	1_2	2			chap.6
HLA_10	HLA	Hawaii		Polynesia	manapohaku	1_2	2			chap.6
HLA_6	HLA	Hawaii		Polynesia	Lanikeka	14	1			chap.6
HLA_7	HLA	Hawaii		Polynesia	palaai	14	1			chap.6
76C610	WBG	Hawaii		Polynesia		1_2	2	1		chap.6
85C393	WBG	Hawaii		Polynesia		1_2	2	1		chap.6
04C6	WBG	Hawaii		Polynesia	lanikeha	1_2	2	1		chap.6
78C463	WBG	Hawaii		Polynesia	Iliahi	6_9	2	78	78A	chap.6
78C464	WBG	Hawaii		Polynesia	Piko	30	1	79		chap.6
04C4	WBG	Hawaii		Polynesia	papa'a kowahi	14	1			chap.6
04C3	WBG	Hawaii		Polynesia	Pala'ai	14	1			chap.6
04C7	WBG	Hawaii		Polynesia	kala	14	1			chap.6
04P8	WBG	Hawaii		Polynesia	Huamoa	14	1			chap.6
89P246	WBG	Hawaii		Polynesia		25	2			chap.6
89P249	WBG	Hawaii		Polynesia		1_2	2			chap.6
78C1121	WBG	Hawaii		Polynesia		1_2	2			chap.6
78C920	WBG	Hawaii		Polynesia	Rapoza (Okinawa?)	6_9	2			chap.6
04C5	WBG	Hawaii		Polynesia	Kalia	14	1			chap.6
05C6	WBG	Hawaii		Polynesia	Mana pohaku	1_2	2			chap.6
78C595	WBG	Hawaii		Polynesia	Nancy Hall	14	1			chap.6
79P636	WBG	Hawaii		Polynesia	Eleele	14	1			chap.6
83C843	WBG	Hawaii		Polynesia		6_9	2			chap.6
89P247	WBG	Hawaii		Polynesia		14	1			chap.6
89P250	WBG	Hawaii		Polynesia		14	1			chap.6
521	Yen	Hawaii		Polynesia	kanaka	14	1			chap.6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
68	Yen	Philippines	Ifugao. BubleiLiqargl	Sout- East Asia	ballaan	1_2	2			chap.6
164	Yen	Philippines	Alab	South-East Asia	butagan	14	1	27		chap.6
53	Yen	Philippines	BayaninanPoitaan	South-East Asia	magyano	14	1			chap.6
57	Yen	Philippines	BayaninanPoitaan	South-East Asia	kayquhin	14	1			chap.6
54	Yen	Philippines	BayaninanPoitaan	South-East Asia	dumaqin	1_2	2			chap.6
10	Yen	Thailandia	Central Plateau	South-East Asia		1_2	2			chap.6
5	Yen	Thailandia	Central Plateau	South-East Asia		1_2	2			chap.6
196	Yen	Timor	Dili	South-East Asia		1_2	2	32		chap.6
185	Yen	Timor	Dili	South-East Asia	mautema	1_2	2			chap.6
193	Yen	Timor	Diquica	South-East Asia		1_2	2	32		chap.6
103	Yen	Philippines	Ifugao, Gohan	South-East Asia	dumaqan	12	1			chap.6
106	Yen	Philippines	Ifugao, Gohan	South-East Asia	bannaqo	1_2	2			chap.6
147	Yen	Philippines	Ifugao, Kiangnan	South-East Asia	tannaya	1_2	2	19	19	chap.6
146	Yen	Philippines	Ifugao, Kiangnan	South-East Asia	talindak	14	1			chap.6
119	Yen	Philippines	Ifugao, Lugu	South-East Asia	kawitan qad hi	1_2	2	11		chap.6
126	Yen	Philippines	Ifugao, Lugu	South-East Asia	qahqahin	ND	ND	12		chap.6
127	Yen	Philippines	Ifugao, Lugu	South-East Asia	daqdaqane	1_2	2	12		chap.6
118	Yen	Philippines	Ifugao, Lugu	South-East Asia	baknat	ND	ND			chap.6
128	Yen	Philippines	Ifugao, Lugu	South-East Asia	qiyambabag	1_2	2			chap.6
121	Yen	Philippines	Ifugao, Lugu	South-East Asia	yawa	1_2	2			chap.6
130	Yen	Philippines	Ifugao, Piwon, Lagawi	South-East Asia	hinyon	1_2	2	11		chap.6
136	Yen	Philippines	Ifugao, Piwon, Lagawi	South-East Asia	gulguldin	ND	ND	17	17	chap.6
141	Yen	Philippines	Ifugao, Piwon, Lagawi	South-East Asia	hobul	1_2	2	19	19	chap.6
135	Yen	Philippines	Ifugao, Piwon, Lagawi	South-East Asia	qontig	1_2	2			chap.6
129	Yen	Philippines	Ifugao, Piwon, Lagawi	South-East Asia	lumnin	1_2	2			chap.6
137	Yen	Philippines	Ifugao, Piwon, Lagawi	South-East Asia	gimpayan	14	1			chap.6
142	Yen	Philippines	Ifugao, Piwon, Lagawi	South-East Asia	bayani	1_2	2			chap.6
70	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia	pukin	1_2	2	11		chap.6
85	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia	kawitan	1_2	2	11		chap.6
81	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia	koyyokoy 2.	1_2	2	17	17	chap.6
73	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia	budqa	1_2	2	77		chap.6
90	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia	qimbaqao	1_2	2	77		chap.6
79	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia		ND	ND			chap.6
86	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia	qiggoltud	14	1			chap.6
71	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia	luglughi	ND	ND			chap.6
74	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia	qinbud	14	1			chap.6
88	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia	pattuki	12	1			chap.6
92	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia	quhohoba	12	1			chap.6
93	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia	tuttolid	ND	ND			chap.6



Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
95	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia	qimbuqdag	14	1			chap.6
98	Yen	Philippines	Ifugao. LiwangMayaoyan	South-East Asia	qubuwan	14	1			chap.6
50	Yen	Philippines	Ifugao. Mountain prov.	South-East Asia	buqagan	1_2	2			chap.6
65	Yen	Philippines	Ifugeo. BubleiLiqargl	South-East Asia	kawitan	1_2	2	11		chap.6
63	Yen	Philippines	Ifugeo. BubleiLiqargl	South-East Asia	ballaan	14	1	27		chap.6
64	Yen	Philippines	Ifugeo. BubleiLiqargl	South-East Asia	lotaan	1_2	2			chap.6
66	Yen	Philippines	Ifugeo. BubleiLiqargl	South-East Asia	baqnat/pukin	ND	ND			chap.6
69	Yen	Philippines	Ifugeo. BubleiLiqargl	South-East Asia		12	1			chap.6
177	Yen	Philippines	Ilocos	South-East Asia	illagan	12	1	28	28	chap.6
179	Yen	Philippines	Ilocos	South-East Asia	ubi	14	1	28	28	chap.6
181	Yen	Philippines	Ilocos	South-East Asia		12	1	28	28	chap.6
173	Yen	Philippines	Ilocos N	South-East Asia	immubi	ND	ND	28	28	chap.6
174	Yen	Philippines	Ilocos N	South-East Asia	baisin	ND	ND	28		chap.6
188	Yen	Timor	Liquica	South-East Asia		ND	ND			chap.6
189	Yen	Timor	Liquica	South-East Asia		6_9	2			chap.6
218	Yen	Philippines	Merauke	South-East Asia		14	1			chap.6
36	Yen	Japan	Okinawa, Ishigaki	South-East Asia	hyakugo	ND	ND	35		chap.6
34	Yen	Japan	Okinawa, Ishigaki	South-East Asia	telma	6_9	2	49		chap.6
37	Yen	Japan	Okinawa, Ishigaki	South-East Asia	sekaichizo	6_9	2	51	51	chap.6
39	Yen	Japan	Okinawa, Ishigaki	South-East Asia	hayainnayo	12	1			chap.6
46	Yen	Japan	Okinawa, Kobama	South-East Asia	sekaichigo	6_9	2	51		chap.6
47	Yen	Japan	Okinawa, Kobama	South-East Asia	baka	6_9	2	51	51	chap.6
40	Yen	Japan	Okinawa, Kobama	South-East Asia	tumaikoru	1_2	2			chap.6
41	Yen	Japan	Okinawa, Kobama	South-East Asia	boko	ND	ND			chap.6
156	Yen	Philippines	Sagada	South-East Asia	qabbaw	1_2	2	11		chap.6
157	Yen	Philippines	Sagada	South-East Asia	qilin	14	1			chap.6
158	Yen	Philippines	Sagada	South-East Asia	kinagayqan	1_2	2			chap.6
160	Yen	Philippines	Sagada	South-East Asia	balo	1_2	2			chap.6
159	Yen	Philippines	Sagada	South-East Asia	qallusiqis	14	1			chap.6
21	Yen	China	South	South-East Asia		1_2	2	35	35	chap.6
22	Yen	China	South	South-East Asia		1_2	2	35	35	chap.6
15	Yen	China	South	South-East Asia		1_2	2			chap.6
17	Yen	China	South	South-East Asia		ND	ND			chap.6
20	Yen	China	South	South-East Asia		14	1			chap.6
440434	CIP	China		South-East Asia		13	1			chap.6
440436	CIP	China		South-East Asia		6_9	2			chap.6
440221	CIP	Japan		South-East Asia		1_2	2			chap.6
440551	CIP	Japan		South-East Asia		14	1			chap.6
440617	CIP	Malaysia		South-East Asia		1_2	2			chap.6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
440620	CIP	Malaysia		South-East Asia		6_9	2			chap.6
440160	CIP	Philippines		South-East Asia		14	1			chap.6
440187	CIP	Philippines		South-East Asia		1_2	2			chap.6
440290	CIP	Philippines		South-East Asia		1_2	2			chap.6
440291	CIP	Philippines		South-East Asia		1_2	2			chap.6
440427	CIP	Philippines		South-East Asia		1_2	2			chap.6
440312	CIP	Taiwan		South-East Asia		1_2	2			chap.6
440409	CIP	Taiwan		South-East Asia		6_9	2			chap.6
440197	CIP	Thailandia		South-East Asia		6_9	2			chap.6
440198	CIP	Thailandia		South-East Asia		1_2	2			chap.6
440334	CIP	Thailandia		South-East Asia		6_9	2			chap.6
440343	CIP	Thailandia		South-East Asia		14	1			chap.6
440348	CIP	Thailandia		South-East Asia		1_2	2			chap.6
440352	CIP	Thailandia		South-East Asia		14	1			chap.6
440353	CIP	Thailandia		South-East Asia		14	1			chap.6
440267	CIP	Vietnam		South-East Asia		14	1			chap.6
440268	CIP	Vietnam		South-East Asia		1_2	2			chap.6
400339	CIP	Brazil	Amazonas	Southern region		1_2	2			chap.3, 5 et 6
420414	CIP	Peru	Amazonas	Southern region	Camote Chancaina	16_23	2			chap.3, 5 et 6
420151	CIP	Peru	Amazonas	Southern region	Camote Potocho	4	1			chap.3, 5 et 6
420386	CIP	Peru	Amazonas	Southern region	Acaro	13	1			chap.3, 5 et 6
420053	CIP	Peru	Ancash	Southern region	Capadito	13	1			chap.3, 5 et 6
420056	CIP	Peru	Ancash	Southern region	Manta China	13	1			chap.3, 5 et 6
420793	CIP	Peru	Ancash	Southern region	Negro	16_23	2			chap.3, 5 et 6
420020	CIP	Peru	Ancash	Southern region	Huarmeyano	14	1			chap.3, 5 et 6
420052	CIP	Peru	Ancash	Southern region	Manzachina	14	1			chap.3, 5 et 6
420362	CIP	Peru	Ancash	Southern region	Culebrero # 2	13	1			chap.3, 5 et 6
422546	CIP	Peru	Ancash	Southern region	Huambacho T1	13	1			chap.3, 5 et 6
420350	CIP	Peru	Arequipa	Southern region	Peludo	14	1			chap.3, 5 et 6
420071	CIP	Peru	Arequipa	Southern region	Milagro	14	1			chap.3, 5 et 6
422536	CIP	Peru	Arequipa	Southern region	Yema de Huevo Arequipeño	19	1			chap.3, 5 et 6
420141	CIP	Peru	Ayacucho	Southern region	Camote Morado	5	1			chap.3, 5 et 6
401042	CIP	Ecuador	Azuay	Southern region	De Sal	13	1			chap.3, 5 et 6
401028	CIP	Ecuador	Azuay	Southern region	Cochinilla	16_23	2			chap.3, 5 et 6
401032	CIP	Ecuador	Azuay	Southern region	Bunuelo	16_23	2			chap.3, 5 et 6
401036	CIP	Ecuador	Azuay	Southern region	Blanco de Dulce	13	1			chap.3, 5 et 6
401021	CIP	Ecuador	Bolivar	Southern region	De Sal	14	1			chap.3, 5 et 6
420027	CIP	Peru	Cajamarca	Southern region	Zapallo	20	1			chap.3, 5 et 6
420025	CIP	Peru	Cajamarca	Southern region	Camote Yuca	19	1			chap.3, 5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
420033	CIP	Peru	Cajamarca	Southern region	ZAPALLO o COSTENO	19	1			chap.3, 5 et 6
420202	CIP	Peru	Cajamarca	Southern region	Chiclayo	8	1			chap.3, 5 et 6
420664	CIP	Peru	Cajamarca	Southern region	Camote Silvestre	14	1			chap.3, 5 et 6
420746	CIP	Peru	Cajamarca	Southern region	Quilikano	14	1			chap.3, 5 et 6
420759	CIP	Peru	Cajamarca	Southern region	Morado	13	1			chap.3, 5 et 6
400939	CIP	Ecuador	Canar	Southern region		14	1			chap.3, 5 et 6
400966	CIP	Ecuador	Carchi	Southern region	Niguilla	13	1			chap.3, 5 et 6
400982	CIP	Ecuador	Carchi	Southern region	Tresmesino	14	1			chap.3, 5 et 6
400968	CIP	Ecuador	Carchi	Southern region	Vera	14	1			chap.3, 5 et 6
609	Yen	Peru	Chicama V	Southern region	pierma de viuda	1_2	2	70		chap.3, 5 et 6
401024	CIP	Ecuador	Chimborazo	Southern region	Guabalqueno	14	1			chap.3, 5 et 6
401533	CIP	Ecuador	Chimborazo	Southern region		6_9	2			chap.3, 5 et 6
615	Yen	Peru	China V	Southern region	porto viejo	ND	ND	34		chap.3, 5 et 6
620(1)	Yen	Peru	China V	Southern region		ND	ND	91	91	chap.3, 5 et 6
622	Yen	Peru	China V	Southern region	porto viejo	14	1			chap.3, 5 et 6
626	Yen	Peru	China V	Southern region	paramonguina	13	1			chap.3, 5 et 6
610	Yen	Peru	Chipillico V	Southern region	otero	14	1			chap.3, 5 et 6
611_1	Yen	Peru	Chipillico V	Southern region	pava blanca	ND	ND			chap.3, 5 et 6
400180	CIP	Bolivia	Chuquisaca	Southern region	Quebradeno	21	1			chap.3, 5 et 6
400171	CIP	Bolivia	Chuquisaca	Southern region	Amarillo	14	1			chap.3, 5 et 6
400176	CIP	Bolivia	Chuquisaca	Southern region	Ogliri o Camote Humedo	14	1			chap.3, 5 et 6
400166	CIP	Bolivia	Cochabamba	Southern region	Blanco	12	1			chap.3, 5 et 6
400162	CIP	Bolivia	Cochabamba	Southern region	Blanco Papa	6_9	2			chap.3, 5 et 6
401566	CIP	Argentina	Corrientes	Southern region	Silvestre Bat x spp?	6_9	2			chap.3, 5 et 6
400484	CIP	Columbia	Cundinamarca	Southern region		6_9	2			chap.3, 5 et 6
420079	CIP	Peru	Cusco	Southern region	Amarillo Cuzqueno	13	1			chap.3, 5 et 6
420334	CIP	Peru	Cusco	Southern region	Zapallo Cuzqueno	13	1			chap.3, 5 et 6
400955	CIP	Ecuador	El Oro	Southern region	Blanco De Sal	13	1			chap.3, 5 et 6
400954	CIP	Ecuador	El Oro	Southern region	Amarillo	13	1			chap.3, 5 et 6
401043	CIP	Ecuador	El Oro	Southern region	Camote Dulce	14	1			chap.3, 5 et 6
691	Yen	Ecuador	Equator	Southern region	paramonguino	14	1			chap.3, 5 et 6
400887	CIP	Ecuador	Esmeralda	Southern region		14	1			chap.3, 5 et 6
692	Yen	Ecuador	Guayaquil	Southern region	paramonguino	14	1			chap.3, 5 et 6
637	Yen	Peru	Huacho	Southern region		13	1			chap.3, 5 et 6
638	Yen	Peru	Huacho	Southern region		13	1			chap.3, 5 et 6
639	Yen	Peru	Huacho	Southern region		19	1			chap.3, 5 et 6
641	Yen	Peru	Huacho	Southern region		14	1			chap.3, 5 et 6
400397	CIP	Columbia	Huila	Southern region	Amarilla	13	1			chap.3, 5 et 6
400387	CIP	Columbia	Huila	Southern region		6_9	2			chap.3, 5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
400389	CIP	Columbia	Huila	Southern region	Huasca	16_23	2			chap.3, 5 et 6
400391	CIP	Columbia	Huila	Southern region		16_23	2			chap.3, 5 et 6
400398	CIP	Columbia	Huila	Southern region	Morada	6_9	2			chap.3, 5 et 6
628	Yen	Peru	Ia Encalada. Trujillo	Southern region	boca de chisco	13	1			chap.3, 5 et 6
400962	CIP	Ecuador	Ibabura	Southern region	Blanco	13	1			chap.3, 5 et 6
420332	CIP	Peru	Ica	Southern region	Chileno de Calidad	14	1			chap.3, 5 et 6
421128	CIP	Peru	Ica	Southern region	Bertita de Humay	14	1			chap.3, 5 et 6
420309	CIP	Peru	Ica	Southern region	Chinchano	14	1			chap.3, 5 et 6
420335	CIP	Peru	Ica	Southern region	Bisambra	14	1			chap.3, 5 et 6
421103	CIP	Peru	Ica	Southern region	Corazon de Huarango Negro	5	1			chap.3, 5 et 6
400974	CIP	Ecuador	Imbabura	Southern region	Camote Papa	14	1			chap.3, 5 et 6
660	Yen	Peru	Julian, Callejeon de Huaylas	Southern region	paramonguino	14	1	67		chap.3, 5 et 6
420368	CIP	Peru	Junin	Southern region	Camote Morado Oscuro	13	1			chap.3, 5 et 6
421024	CIP	Peru	La Libertad	Southern region	Amarillo	13	1			chap.3, 5 et 6
420039	CIP	Peru	La Libertad	Southern region	Esperma	14	1			chap.3, 5 et 6
420564	CIP	Peru	La Libertad	Southern region	Sanpedrano	14	1			chap.3, 5 et 6
421027	CIP	Peru	La Libertad	Southern region	Camote Yuca	14	1			chap.3, 5 et 6
421113	CIP	Peru	La Libertad	Southern region	Amarillo de Trujillo	14	1			chap.3, 5 et 6
420047	CIP	Peru	Lambayeque	Southern region	Estrella	14	1			chap.3, 5 et 6
420338	CIP	Peru	Lambayeque	Southern region	Amarillo de Chiclayo	13	1			chap.3, 5 et 6
420031	CIP	Peru	Lima	Southern region	Cascajo Morado	11	1			chap.3, 5 et 6
420304	CIP	Peru	Lima	Southern region	De Yuca	13	1			chap.3, 5 et 6
421091	CIP	Peru	Lima	Southern region	San Juanino	13	1			chap.3, 5 et 6
420361	CIP	Peru	Lima	Southern region	Pativilcano	1_2	2			chap.3, 5 et 6
420029	CIP	Peru	Lima	Southern region	Papel	14	1			chap.3, 5 et 6
420064	CIP	Peru	Lima	Southern region	Paramonguino	14	1			chap.3, 5 et 6
420066	CIP	Peru	Lima	Southern region	Ihuanco	14	1			chap.3, 5 et 6
420301	CIP	Peru	Lima	Southern region	Italiano Plomo	19	1			chap.3, 5 et 6
420303	CIP	Peru	Lima	Southern region	Yema de Huevo	14	1			chap.3, 5 et 6
420317	CIP	Peru	Lima	Southern region	Morado De Huaura	14	1			chap.3, 5 et 6
420321	CIP	Peru	Lima	Southern region	Morado de Canete	14	1			chap.3, 5 et 6
420324	CIP	Peru	Lima	Southern region	Amarillo de Rinconada	13	1			chap.3, 5 et 6
420337	CIP	Peru	Lima	Southern region	Blanco De Huaura	13	1			chap.3, 5 et 6
421106	CIP	Peru	Lima	Southern region	Japones Portugues	13	1			chap.3, 5 et 6
644	Yen	Peru	Lima	Southern region		13	1	93	93	chap.3, 5 et 6
400949	CIP	Ecuador	Loja	Southern region		8	1			chap.3, 5 et 6
420652	CIP	Peru	Loreto	Southern region	Camote Silvestre	8	1			chap.3, 5 et 6
400958	CIP	Ecuador	Los Rios	Southern region	Camote	8	1			chap.3, 5 et 6
401045	CIP	Ecuador	Los Rios	Southern region	Manaba de Dulce	14	1			chap.3, 5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
400002	CIP	Ecuador	Manabi	Southern region	Morado	14	1			chap.3, 5 et 6
400223	CIP	Brazil	Mato Grosso	Southern region		1_2	2			chap.3, 5 et 6
400236	CIP	Brazil	Mato Grosso	Southern region		16_23	2			chap.3, 5 et 6
400248	CIP	Brazil	Mato Grosso	Southern region	Branca	14	1			chap.3, 5 et 6
400032	CIP	Paraguay	Misiones	Southern region	Yety Aba	16_23	2			chap.3, 5 et 6
630_A	Yen	Peru	Moche V	Southern region	blanco esperma	13	1	91	91	chap.3, 5 et 6
420213	CIP	Peru	Moquegua	Southern region		16_23	2			chap.3, 5 et 6
420051	CIP	Peru	Moquegua	Southern region	Chinchano	14	1			chap.3, 5 et 6
400941	CIP	Ecuador	Morona	Southern region	Camote de Dulce	13	1			chap.3, 5 et 6
400904	CIP	Ecuador	Napo	Southern region	Dulce	16_23	2	56		chap.3, 5 et 6
401525	CIP	Ecuador	Napo	Southern region		16_23	2	56		chap.3, 5 et 6
400899	CIP	Ecuador	Napo	Southern region	De Dulce	15	1			chap.3, 5 et 6
400906	CIP	Ecuador	Napo	Southern region	Jibaro	15	1			chap.3, 5 et 6
400908	CIP	Ecuador	Napo	Southern region	Blanco Dulce	15	1			chap.3, 5 et 6
400420	CIP	Columbia	Narino	Southern region		13	1	105		chap.3, 5 et 6
400421	CIP	Columbia	Narino	Southern region	Corre Huasca	ND	ND	105		chap.3, 5 et 6
400436	CIP	Columbia	Narino	Southern region	Froilana	13	1			chap.3, 5 et 6
400459	CIP	Columbia	Narino	Southern region	Morada Pequena	13	1			chap.3, 5 et 6
400443	CIP	Columbia	Narino	Southern region	Camota	6_9	2			chap.3, 5 et 6
400458	CIP	Columbia	Narino	Southern region	Blanca de Sal	4	1			chap.3, 5 et 6
400430	CIP	Columbia	Narino	Southern region	Huilenza	6_9	2			chap.3, 5 et 6
400450	CIP	Columbia	Narino	Southern region	Bogotana	6_9	2			chap.3, 5 et 6
400211	CIP	Bolivia	Nor Yungas	Southern region	Blanco	6_9	2			chap.3, 5 et 6
400212	CIP	Bolivia	Nor Yungas	Southern region	Amarillo	14	1			chap.3, 5 et 6
421126	CIP	Peru	Pasco	Southern region	Camote De Oxapampa	13	1			chap.3, 5 et 6
400929	CIP	Ecuador	Pastaza	Southern region	De Dulce	16_23	2			chap.3, 5 et 6
400031	CIP	Paraguay	Pdte Hayes	Southern region	Yety Manduvi	14	1			chap.3, 5 et 6
400987	CIP	Ecuador	Pichincha	Southern region	Guayaco	1_2	2	57		chap.3, 5 et 6
401005	CIP	Ecuador	Pichincha	Southern region	Jumbo Dulce o Tanda	13	1			chap.3, 5 et 6
400991	CIP	Ecuador	Pichincha	Southern region	Polvarron	6_9	2			chap.3, 5 et 6
420607	CIP	Peru	Piura	Southern region	Trujillano	11	1			chap.3, 5 et 6
420269	CIP	Peru	Piura	Southern region	Napa	14	1			chap.3, 5 et 6
420356	CIP	Peru	Piura	Southern region	Bejuco	14	1			chap.3, 5 et 6
420580	CIP	Peru	Piura	Southern region	Blanco	13	1			chap.3, 5 et 6
420617	CIP	Peru	Piura	Southern region	Chilpo Blanco	13	1			chap.3, 5 et 6
600	Yen	Peru	Piura V	Southern region		1_2	2	35	35	chap.3, 5 et 6
420201	CIP	Peru	Puno	Southern region		14	1			chap.3, 5 et 6
421007	CIP	Peru	Puno	Southern region	Papa Camote	14	1			chap.3, 5 et 6
400467	CIP	Columbia	Quindio	Southern region	Rojo	6_9	2			chap.3, 5 et 6

Number	<i>Ex situ</i> Collection	Country/archipelago	Island/province/locality	Region	Local name	H	Cp lineage	MCL	MLG	Reference
400479	CIP	Columbia	Quindio	Southern region		6_9	2			chap.3, 5 et 6
666	Yen	Peru	Rocuay Ancash	Southern region	paramonguino	14	1	67	67	chap.3, 5 et 6
665_2	Yen	Peru	Rocuay Ancash	Southern region	paramonguino	6_9	2	72		chap.3, 5 et 6
400283	CIP	Brazil	Rondonia	Southern region	Roxa	1_2	2			chap.3, 5 et 6
400312	CIP	Brazil	Rondonia	Southern region	Vermelha	14	1			chap.3, 5 et 6
695	Yen	Ecuador	San Horca	Southern region		13	1	59		chap.3, 5 et 6
420249	CIP	Peru	San Martin	Southern region	Norteno	14	1			chap.3, 5 et 6
400044	CIP	Argentina	Stgo Del Estero	Southern region	Criolla Colorada	14	1	2	2	chap.3, 5 et 6
421157	CIP	Peru	Tacna	Southern region	Supano Negro	13	1			chap.3, 5 et 6
400220	CIP	Bolivia	Tarija	Southern region	Colorada	14	1			chap.3, 5 et 6
400483	CIP	Columbia	Tolima	Southern region	Batata	16_23	2			chap.3, 5 et 6
400142	CIP	Argentina	Tucumam	Southern region	Rojo Blanco	14	1			chap.3, 5 et 6
420325	CIP	Peru	Tumbes	Southern region	Locro	13	1			chap.3, 5 et 6
420634	CIP	Peru	Tumbes	Southern region	Rosa Amarillo	14	1			chap.3, 5 et 6
401017	CIP	Ecuador	Tungurahua	Southern region		13	1	59		chap.3, 5 et 6
400476	CIP	Columbia	Valle De Cauca	Southern region	Batata Roja	6_9	2			chap.3, 5 et 6
401536	CIP	Ecuador	Zamora	Southern region		4	1			chap.3, 5 et 6
421100	CIP	Peru		Southern region		ND	ND			chap.3, 5 et 6

**Annexe 1: Données passeport des échantillons analysés dans les différents chapitres de cette étude.**

b) Echantillons du Chapitre 2

Sample number	Species	collection	ploidy	Country	haplo _ITS	haplo _IGS	cp SSR group	Struct. Cluster	K means cluster	Latitude	Longitude
400044	<i>I. batatas</i>	CIP	6x	Argentina			1	1	1	-27.83	-63.66
400180	<i>I. batatas</i>	CIP	6x	Bolivia		19	1	1	1	-18.83	-64.83
400166	<i>I. batatas</i>	CIP	6x	Bolivia			1	1	1	-17	-65.66
401559	<i>I. batatas</i>	CIP	6x	Colombia		6	2	5	2	8.08	-73.16
400493	<i>I. batatas</i>	CIP	6x	Colombia	12	12	2	4	4	11.25	-73.61
400512	<i>I. batatas</i>	CIP	6x	Colombia		12	2	5	3	8.35	-73.51
400436	<i>I. batatas</i>	CIP	6x	Colombia	6	19	1	1	1	1.6	-76.91
400377	<i>I. batatas</i>	CIP	6x	Colombia			2	4	4	9.91	-75.33
400382	<i>I. batatas</i>	CIP	6x	Colombia				4	4	11.5	-72.83
400397	<i>I. batatas</i>	CIP	6x	Colombia			1	1	1	2.41	-75.86
400420	<i>I. batatas</i>	CIP	6x	Colombia			1	1	1	1.5	-77.33
400421	<i>I. batatas</i>	CIP	6x	Colombia				1	1	1.3	-77.16
400443	<i>I. batatas</i>	CIP	6x	Colombia			2	1	1	1.46	-77.58
400458	<i>I. batatas</i>	CIP	6x	Colombia			1	1	1	1.16	-77.53
400459	<i>I. batatas</i>	CIP	6x	Colombia			1	1	1	1.25	-77.25
400509	<i>I. batatas</i>	CIP	6x	Colombia			1	3	3	10.41	-73.75
400515	<i>I. batatas</i>	CIP	6x	Colombia			2	5	1	8.08	-73.05
400517	<i>I. batatas</i>	CIP	6x	Colombia			1	3	3	8.06	-73
400518	<i>I. batatas</i>	CIP	6x	Colombia			1	3	3	8.33	-73.43
401551	<i>I. batatas</i>	CIP	6x	Colombia			1	1	1	2.36	-77.66
400551	<i>I. batatas</i>	CIP	6x	Costa Rica			1	5	3	10.4176	-84.430847
400578	<i>I. batatas</i>	CIP	6x	Cuba			2	3	1	21.8819	-78.643799
400584	<i>I. batatas</i>	CIP	6x	Cuba			1	3	3	ND	ND
400602	<i>I. batatas</i>	CIP	6x	Cuba			2	3	3	ND	ND
400608	<i>I. batatas</i>	CIP	6x	Cuba			1	3	3	ND	ND
400034	<i>I. batatas</i>	CIP	6x	Dominican Republic			2	3	3	ND	ND
400786	<i>I. batatas</i>	CIP	6x	Dominican Republic			1	3	3	ND	ND
400815	<i>I. batatas</i>	CIP	6x	Dominican Republic				3	3	18.66	-70.75
400817	<i>I. batatas</i>	CIP	6x	Dominican Republic			1	3	3	18.66	-70.75
400825	<i>I. batatas</i>	CIP	6x	Dominican Republic			2	3	3	19.6	-70.5
400828	<i>I. batatas</i>	CIP	6x	Dominican Republic			1	3	3	19.56	-70.53
400830	<i>I. batatas</i>	CIP	6x	Dominican Republic			1	3	3	19.16	-70.5
400865	<i>I. batatas</i>	CIP	6x	Dominican Republic			2	3	3	-0.83	-77.5
401541	<i>I. batatas</i>	CIP	6x	Dominican Republic			2	3	3	18.66	-70.75
401525	<i>I. batatas</i>	CIP	6x	Ecuador	8	12	2	1	1	-0.16	-76.91
400908	<i>I. batatas</i>	CIP	6x	Ecuador		19	1	5	1	-0.83	-76.5
400968	<i>I. batatas</i>	CIP	6x	Ecuador		19	1	1	1	0.58	-71.5
401021	<i>I. batatas</i>	CIP	6x	Ecuador	41	19	1	1	1	-1.66	-79.08
401017	<i>I. batatas</i>	CIP	6x	Ecuador		20	1	1	1	-1.41	-78.41
400904	<i>I. batatas</i>	CIP	6x	Ecuador			2	1	1	-0.16	-76.91
400929	<i>I. batatas</i>	CIP	6x	Ecuador			2	1	1	-1.46	-78.13
400941	<i>I. batatas</i>	CIP	6x	Ecuador			1	1	1	-3.5	-78.58
400954	<i>I. batatas</i>	CIP	6x	Ecuador			1	1	1	-3.63	-79.75
400955	<i>I. batatas</i>	CIP	6x	Ecuador			1	1	1	-3.66	-79.75
400962	<i>I. batatas</i>	CIP	6x	Ecuador			1	1	1	-0.66	-78.16
400966	<i>I. batatas</i>	CIP	6x	Ecuador			1	1	1	0.5	-78.16
400982	<i>I. batatas</i>	CIP	6x	Ecuador			1	5	1	0.5	-78.08
400987	<i>I. batatas</i>	CIP	6x	Ecuador			2	3	1	-0.16	-78.33
401005	<i>I. batatas</i>	CIP	6x	Ecuador			1	4	4	1.16	-78.41
401024	<i>I. batatas</i>	CIP	6x	Ecuador			1	1	1	-2.16	-78.91
401028	<i>I. batatas</i>	CIP	6x	Ecuador			2	5	4	-2.75	-78.66
401032	<i>I. batatas</i>	CIP	6x	Ecuador			2	5	4	-3.16	-79.08
401042	<i>I. batatas</i>	CIP	6x	Ecuador			1	1	1	-3.33	-79.41
401533	<i>I. batatas</i>	CIP	6x	Ecuador			2	1	1	-2.33	-79.16
401079	<i>I. batatas</i>	CIP	6x	Guatemala	13	12	2	4	4	14.53	-90.76
401083	<i>I. batatas</i>	CIP	6x	Guatemala		12	2	4	4	14.4166	-90.25
401094	<i>I. batatas</i>	CIP	6x	Guatemala	11	12	2	5	3	16.95	-89.48
401055	<i>I. batatas</i>	CIP	6x	Guatemala			1	5	1	14.28	-89.81
401101	<i>I. batatas</i>	CIP	6x	Guatemala			2	3	1	15.46	-88.81
401114	<i>I. batatas</i>	CIP	6x	Guatemala			1	5	2	15.31	-89.9



401119	<i>I. batatas</i>	CIP	6x	Guatemala			1	5	1	15.05	-90.23
401136	<i>I. batatas</i>	CIP	6x	Guatemala			2	4	4	14.51	-91.25
401587	<i>I. batatas</i>	CIP	6x	Guatemala			1	3	3	15.3	-91.06
401183	<i>I. batatas</i>	CIP	6x	Jamaica	12		2	5	3	18.08	-77.16
400040	<i>I. batatas</i>	CIP	6x	Jamaica			2	5	3	18	-77.08
401170	<i>I. batatas</i>	CIP	6x	Jamaica			2	5	3	18.33	-77.16
401172	<i>I. batatas</i>	CIP	6x	Jamaica			2	5	3	18.58	-77.25
401200	<i>I. batatas</i>	CIP	6x	Jamaica			1	3	3	17.83	-77.03
400010	<i>I. batatas</i>	CIP	6x	Mexico	10	12	2	3	3	ND	ND
400009	<i>I. batatas</i>	CIP	6x	Mexico			1	3	3	ND	ND
401207	<i>I. batatas</i>	CIP	6x	Mexico			2	5	1	22.6748	-101.79199
401208	<i>I. batatas</i>	CIP	6x	Mexico			2	5	1	ND	ND
401210	<i>I. batatas</i>	CIP	6x	Mexico				5	1	ND	ND
401213	<i>I. batatas</i>	CIP	6x	Mexico			2	3	3	19.0666	-96.9833
401214	<i>I. batatas</i>	CIP	6x	Mexico			2	3	3	ND	ND
401215	<i>I. batatas</i>	CIP	6x	Mexico			2	3	3	21.3166	-97.5
401475	<i>I. batatas</i>	CIP	6x	Mexico			1	5	1	ND	ND
441724	<i>I. batatas</i>	CIP	6x	Mexico			2	5	4	ND	ND
401269	<i>I. batatas</i>	CIP	6x	Panama	12		2	4	4	8.75	-78.46
401272	<i>I. batatas</i>	CIP	6x	Panama	12		2	4	4	8.13	-78.03
401236	<i>I. batatas</i>	CIP	6x	Panama			1	5	4	8.58	-79.96
401238	<i>I. batatas</i>	CIP	6x	Panama			2	5	1	7.75	-80.71
401248	<i>I. batatas</i>	CIP	6x	Panama			2	1	1	8.36	-81.23
401251	<i>I. batatas</i>	CIP	6x	Panama			2	4	4	8.1	-81.03
401259	<i>I. batatas</i>	CIP	6x	Panama			2	4	4	8.16	-80.43
401264	<i>I. batatas</i>	CIP	6x	Panama			2	4	4	9.08	-78.91
401267	<i>I. batatas</i>	CIP	6x	Panama			2	4	4	8.8	-78.5
401277	<i>I. batatas</i>	CIP	6x	Panama			1	3	3	9.13	-79.03
401280	<i>I. batatas</i>	CIP	6x	Panama			1	5	1	8.9	-79.73
420033	<i>I. batatas</i>	CIP	6x	Peru	12		2	1	1	-7.55	-78.5
420027	<i>I. batatas</i>	CIP	6x	Peru	19		1	1	1	-7.75	-78.16
420324	<i>I. batatas</i>	CIP	6x	Peru	20		1	1	1	-13	-76.5
420845	<i>I. batatas</i>	CIP	6x	Peru	20		1	1	1	-6.33	-76.53
421128	<i>I. batatas</i>	CIP	6x	Peru	20		1	1	1	-14	-76
420031	<i>I. batatas</i>	CIP	6x	Peru			1	1	1	-11.27	-76.81
420053	<i>I. batatas</i>	CIP	6x	Peru			1	1	1	-9.25	-77.83
420056	<i>I. batatas</i>	CIP	6x	Peru			1	1	1	-9.16	-77.83
420079	<i>I. batatas</i>	CIP	6x	Peru			1	5	1	-13	-73
420213	<i>I. batatas</i>	CIP	6x	Peru			2	1	1	-10.25	-70.96
420304	<i>I. batatas</i>	CIP	6x	Peru			1	5	1	-12	-77
420309	<i>I. batatas</i>	CIP	6x	Peru			1	1	1	-15	-75.5
420325	<i>I. batatas</i>	CIP	6x	Peru			1	1	1	-3.5	-80.5
420332	<i>I. batatas</i>	CIP	6x	Peru			1	1	1	-14	-76.5
420350	<i>I. batatas</i>	CIP	6x	Peru			1	1	1	-15.5	-74.5
420361	<i>I. batatas</i>	CIP	6x	Peru			2	5	3	-11	-78
420368	<i>I. batatas</i>	CIP	6x	Peru			1	1	1	-11.18	-74.68
420414	<i>I. batatas</i>	CIP	6x	Peru			2	1	1	-6.83	-78.01
420607	<i>I. batatas</i>	CIP	6x	Peru			1	5	1	-4.5	-80
420634	<i>I. batatas</i>	CIP	6x	Peru			1	5	3	-3.83	-80.33
420793	<i>I. batatas</i>	CIP	6x	Peru			2	1	1	-9.5	-77.83
421024	<i>I. batatas</i>	CIP	6x	Peru			1	1	1	-8.51	-78.8
421027	<i>I. batatas</i>	CIP	6x	Peru			1	1	1	-7.7	-79.41
421091	<i>I. batatas</i>	CIP	6x	Peru			1	1	1	-12	-77
421100	<i>I. batatas</i>	CIP	6x	Peru				1	1	-14	-76
400031	<i>I. batatas</i>	CIP	6x	Puerto Rico			1	5	1	-24.08	-59.16
400032	<i>I. batatas</i>	CIP	6x	Puerto Rico			2	1	1	-26.81	-56.9
400024	<i>I. batatas</i>	CIP	6x	Saint Vincent	7	12	2	5	3	13.264	-61.194305
400025	<i>I. batatas</i>	CIP	6x	Saint Vincent			1	5	1	13.264	-61.194305
401394	<i>I. batatas</i>	CIP	6x	Venezuela	9	12	2	4	4	9.16	-70.16
401446	<i>I. batatas</i>	CIP	6x	Venezuela		12		3	3	10.08	-63.66
401378	<i>I. batatas</i>	CIP	6x	Venezuela			2	3	3	9.66	-68.66
401381	<i>I. batatas</i>	CIP	6x	Venezuela			2	5	1	9.75	-69.66
401392	<i>I. batatas</i>	CIP	6x	Venezuela			2	5	3	9.16	-70.16

401396	<i>I. batatas</i>	CIP	6x	Venezuela			2	3	3	8.66	-71.75
401399	<i>I. batatas</i>	CIP	6x	Venezuela			1	3	3	8.33	-71.83
401403	<i>I. batatas</i>	CIP	6x	Venezuela			2	5	4	8.58	-71.08
401430	<i>I. batatas</i>	CIP	6x	Venezuela			2	3	3	10.1	-66.6
401431	<i>I. batatas</i>	CIP	6x	Venezuela			2	5	3	9.4	-67.56
401436	<i>I. batatas</i>	CIP	6x	Venezuela			2	5	3	10.26	-63.83
401437	<i>I. batatas</i>	CIP	6x	Venezuela			1	3	3	10.25	-63.8
401445	<i>I. batatas</i>	CIP	6x	Venezuela			1	3	3	10.08	-63.66
401451	<i>I. batatas</i>	CIP	6x	Venezuela			1	3	3	9.86	-63.2
401453	<i>I. batatas</i>	CIP	6x	Venezuela			2	5	3	10.03	-63.63
401454	<i>I. batatas</i>	CIP	6x	Venezuela			1	3	3	10.06	-63.63
401455	<i>I. batatas</i>	CIP	6x	Venezuela			2	3	3	10.16	-63.5
401464	<i>I. batatas</i>	CIP	6x	Venezuela			2	4	4	9.16	-70.16
401500	<i>I. batatas</i>	CIP	6x	Venezuela			1	5	4	8.58	-71.16
401501	<i>I. batatas</i>	CIP	6x	Venezuela			2	4	4	8.83	-70.5
401560	<i>I. batatas</i>	CIP	6x	Venezuela			1	3	3	11.1	-63.83
460104_1	<i>Ipomoea leucantha</i>	CIP	2x	Colombia	39	3				10.43	-73.2
460104_2	<i>Ipomoea leucantha</i>	CIP	2x	Colombia	39	3				10.43	-73.2
460002_1	<i>Ipomoea leucantha</i>	CIP	2x	Ecuador		3				-3.75	-79.91
460002_2	<i>Ipomoea leucantha</i>	CIP	2x	Ecuador		3				-3.75	-79.91
460003_1	<i>Ipomoea leucantha</i>	CIP	2x	Peru		8				-8.66	-74.33
460003_2	<i>Ipomoea leucantha</i>	CIP	2x	Peru		8				-8.66	-74.33
403261_1	<i>Ipomoea sp.</i>	CIP	4x	Colombia		1		2	2	4.33	-75.61
403261_2	<i>Ipomoea sp.</i>	CIP	4x	Colombia		1		2	2	4.33	-75.61
403248_2	<i>Ipomoea sp.</i>	CIP	4x	Colombia		12	2	2	2	3.66	-76.5
403248_3	<i>Ipomoea sp.</i>	CIP	4x	Colombia		12	2	5	4	3.66	-76.5
403261_3	<i>Ipomoea sp.</i>	CIP	4x	Colombia		12	2	2	2	4.33	-75.61
403270_1	<i>Ipomoea sp.</i>	CIP	4x	Colombia		12	2	2	2	4.58	-75.66
403270_2	<i>Ipomoea sp.</i>	CIP	4x	Colombia		12	2	2	2	4.58	-75.66
K500_1	<i>Ipomoea sp.</i>	NIAS	4x	Colombia		12	2	2	2	3.46065	-76.523895
K500_2	<i>Ipomoea sp.</i>	NIAS	4x	Colombia		12	2	2	2	3.46065	-76.523895
K500_4	<i>Ipomoea sp.</i>	NIAS	4x	Colombia		12	2	2	2	3.46065	-76.523895
K300_6	<i>Ipomoea sp.</i>	NIAS	4x	Ecuador	16	12	2	5	2	0.9338	-79.667358
K300_3	<i>Ipomoea sp.</i>	NIAS	4x	Ecuador		13				0.9338	-79.667358
403552_1	<i>Ipomoea sp.</i>	CIP	4x	Ecuador		19	1	5	1	0.86	-79.16
403552_2	<i>Ipomoea sp.</i>	CIP	4x	Ecuador		19	1	5	1	0.86	-79.16
K300_7	<i>Ipomoea sp.</i>	NIAS	4x	Ecuador		19	1	5	1	0.9338	-79.667358
460029_1	<i>Ipomoea sp.</i>	CIP	2x	Ecuador		21	1	2	2	-3.33	-79.91
460029_2	<i>Ipomoea sp.</i>	CIP	2x	Ecuador		21				-3.33	-79.91
460029_3	<i>Ipomoea sp.</i>	CIP	2x	Ecuador		21	1	2	2	-3.33	-79.91
K300_5	<i>Ipomoea sp.</i>	NIAS	4x	Ecuador	2		NA	5	2	0.9338	-79.667358
AL_2145(1	<i>Ipomoea sp.</i>	NIAS	4x	Guatemala		12	2	4	4	ND	ND
AL_2156(1	<i>Ipomoea sp.</i>	NIAS	4x	Guatemala	14	12	2	5	2	14.6221	-90.532837
AL_2232(3	<i>Ipomoea sp.</i>	NIAS	4x	Guatemala	42	12	2	5	4	15.5595	-88.967285
CAL_2192_	<i>Ipomoea sp.</i>	NIAS	4x	Guatemala	15					15.1153	-89.36142
403959_2	<i>Ipomoea sp.</i>	CIP	4x	Mexico		12	2	4	4	16.91	-92.16
CAL_2262_	<i>Ipomoea sp.</i>	NIAS	4x	Mexico		12	2	4	4	17.5516	-92.953434
K123_1	<i>Ipomoea sp.</i>	NIAS	6x	Mexico		12	2	4	4	18.8493	-97.098885
K177_2	<i>Ipomoea sp.</i>	NIAS	6x	Mexico		12	2	4	4	19.5443	-96.920013
K222_2	<i>Ipomoea sp.</i>	NIAS	3x	Mexico		12	2	5	2	ND	ND
403959_1	<i>Ipomoea sp.</i>	CIP	4x	Mexico		19	1	4	4	16.91	-92.16
403959_3	<i>Ipomoea sp.</i>	CIP	4x	Mexico		20	1	4	4	16.91	-92.16
K123_10	<i>Ipomoea sp.</i>	NIAS	6x	Mexico				4	4	18.8493	-97.098885
K123_11	<i>Ipomoea sp.</i>	NIAS	6x	Mexico				4	4	18.8493	-97.098885
K123_12	<i>Ipomoea sp.</i>	NIAS	6x	Mexico				4	4	18.8493	-97.098885
K123_14	<i>Ipomoea sp.</i>	NIAS	6x	Mexico				4	4	18.8493	-97.098885
K123_18	<i>Ipomoea sp.</i>	NIAS	6x	Mexico				4	4	18.8493	-97.098885
K123_20	<i>Ipomoea sp.</i>	NIAS	6x	Mexico				4	4	18.8493	-97.098885
K123_4	<i>Ipomoea sp.</i>	NIAS	6x	Mexico				4	4	18.8493	-97.098885
K123_5	<i>Ipomoea sp.</i>	NIAS	6x	Mexico				4	4	18.8493	-97.098885
K123_8	<i>Ipomoea sp.</i>	NIAS	6x	Mexico				4	4	18.8493	-97.098885
K123_9	<i>Ipomoea sp.</i>	NIAS	6x	Mexico				4	4	18.8493	-97.098885
460824_1	<i>Ipomoea tabasacana</i>	CIP	4x	Mexico	43	12	2	4	4	17.58	-91.25

460824_2	<i>pomoea tabasacana</i>	CIP	4x	Mexico		12	2	5	4	17.58	-91.25
460531_1	<i>Ipomoea tiliacea</i>	CIP	4x	Cuba	1	13				22.25	-80
460531_2	<i>Ipomoea tiliacea</i>	CIP	4x	Cuba	1	13				22.25	-80
460070_1	<i>Ipomoea tiliacea</i>	CIP	4x	Mexico		3				17.83	-93.33
K134_1	<i>Ipomoea tiliacea</i>	NIAS	4x	Mexico		11				ND	ND
K270_8	<i>Ipomoea tiliacea</i>	NIAS	4x	Mexico		13				ND	ND
K134_2	<i>Ipomoea tiliacea</i>	NIAS	4x	Mexico						ND	ND
K270_1	<i>Ipomoea tiliacea</i>	NIAS	4x	Mexico						ND	ND
K270_2	<i>Ipomoea tiliacea</i>	NIAS	4x	Mexico						ND	ND
K270_5	<i>Ipomoea tiliacea</i>	NIAS	4x	Mexico						ND	ND
460551_1	<i>Ipomoea trifida</i>	CIP	2x	Cuba	23					23	-82.5
460551_2	<i>Ipomoea trifida</i>	CIP	2x	Cuba	32					23	-82.5
460108_1	<i>Ipomoea trifida</i>	CIP	2x	Colombia	30	9		2	2	10.66	-74.96
460108_2	<i>Ipomoea trifida</i>	CIP	2x	Colombia		9		2	2	10.66	-74.96
460137_1	<i>Ipomoea trifida</i>	CIP	2x	Colombia		9		2	2	10.25	-74.08
460137_3	<i>Ipomoea trifida</i>	CIP	2x	Colombia		9		2	2	10.25	-74.08
JA0105_2	<i>Ipomoea trifida</i>	USDA	2x	Costa Rica		2		2	2	10.8252	-85.638428
JA0115_2	<i>Ipomoea trifida</i>	USDA	2x	Costa Rica	33	2		2	2	10.8717	-85.76
JA 0100_1	<i>Ipomoea trifida</i>	USDA	2x	Costa Rica		9				10.8252	-85.638428
JA 0115_3	<i>Ipomoea trifida</i>	USDA	2x	Costa Rica		9				10.4797	-85.08
JA0100_3	<i>Ipomoea trifida</i>	USDA	2x	Costa Rica		9		2	2	10.8252	-85.638428
JA0100_4	<i>Ipomoea trifida</i>	USDA	2x	Costa Rica		9		2	2	10.8252	-85.638428
JA0105_1	<i>Ipomoea trifida</i>	USDA	2x	Costa Rica		9		2	2	10.8252	-85.638428
460763_11	<i>Ipomoea trifida</i>	CIP	2x	Guatemala	20	7		2	2	14.21	-90.33
460266	<i>Ipomoea trifida</i>	CIP	2x	Guatemala		9		2	2	14.46	-90.63
460262_5	<i>Ipomoea trifida</i>	CIP	2x	Guatemala		9		2	2	14.31	-90.43
460262_9	<i>Ipomoea trifida</i>	CIP	2x	Guatemala	33	9		2	2	14.31	-90.43
460274_5	<i>Ipomoea trifida</i>	CIP	2x	Guatemala		9		2	2	15.73	-89.16
460274_6	<i>Ipomoea trifida</i>	CIP	2x	Guatemala		9		2	2	15.73	-89.16
460736_2	<i>Ipomoea trifida</i>	CIP	2x	Guatemala		9		2	2	14.66	-92.16
460741_10	<i>Ipomoea trifida</i>	CIP	2x	Guatemala	33	9		2	2	14.3	-90.96
460745_3	<i>Ipomoea trifida</i>	CIP	2x	Guatemala	33	9		2	2	14.25	-90.7
460763_10	<i>Ipomoea trifida</i>	CIP	2x	Guatemala	33	9		2	2	14.21	-90.33
460741_1	<i>Ipomoea trifida</i>	CIP	2x	Guatemala	33	10		2	2	14.3	-90.96
460745_4	<i>Ipomoea trifida</i>	CIP	2x	Guatemala		14		2	2	14.25	-90.7
460262_6	<i>Ipomoea trifida</i>	CIP	2x	Guatemala		15		5	2	14.31	-90.43
Grif 6207_1	<i>Ipomoea trifida</i>	USDA	2x	Mexico		4				ND	ND
Grif 6197_3	<i>Ipomoea trifida</i>	USDA	2x	Mexico		5				16.21	-98.16
Grif 6197_5	<i>Ipomoea trifida</i>	USDA	2x	Mexico		9				16.21	-98.16
Grif 6198_6	<i>Ipomoea trifida</i>	USDA	2x	Mexico		9				16.1	-93.15
Grif 6204_2	<i>Ipomoea trifida</i>	USDA	2x	Mexico		9				ND	ND
460371	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	11.88	-85.95
460375_7	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	11.88	-86.08
460378_1	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	11.43	-85.86
460378_3	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	11.43	-85.86
460379_1	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	11.96	-85.56
460379_10	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	11.96	-85.56
460379_6	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	11.96	-85.56
460380_2	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	11.96	-85.56
460385_18	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	12.33	-86.45
460385_19	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	12.33	-86.45
460387_2	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	12.36	-86.66
460387_4	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	12.36	-86.66
460388_11	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	12.81	-86.58
460388_2	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	12.81	-86.58
460390_10	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	13.35	-86.43
460390_2	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	24	9		2	2	13.3665	-86.50829
460403_1	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	33	9		2	2	12.76	-85.4
460403_2	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	33	9		2	2	12.76	-85.4
460403_3	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9		2	2	12.76	-85.4
460406_1	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	22	9		2	2	12.08	-85.33
460406_19	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	33	9		2	2	12.08	-85.33
460407_1	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	33	9		2	2	12.53	-85.76

460407_8	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	19	9	2	2	12.53	-85.76
460408_13	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	28	9	2	2	12.08	-85.33
460408_14	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		9	2	2	12.08	-85.33
460408_2	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	17	9	2	2	12.08	-85.33
460410_2	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	17	9	2	2	12.01	-85.16
460410_3	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	31	9	2	2	12.01	-85.16
460410_5	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	33	9	2	2	12.01	-85.16
460411_1	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	33	9	2	2	12.03	-85.16
460411_13	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	33	9	2	2	12.03	-85.16
460424_1	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	33	9	2	2	12.4667	-84.38333
460424_13	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	27	9	2	2	12.4667	-84.38333
460429_1	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	25	9	2	2	12.15	-86.26
460429_3	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	33	9	2	2	12.15	-86.26
460431_1	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	22	9	2	2	12.15	-86.26
460431_6	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	21	9	2	2	12.15	-86.26
460431_7	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	33	9	2	2	12.15	-86.26
460387_3	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua		14	2	2	12.36	-86.66
460410_6	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	29	16	2	2	12.01	-85.16
460429_2	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	33	17	2	2	12.15	-86.26
460435_11	<i>Ipomoea trifida</i>	CIP	2x	Nicaragua	26		2	2	12.15	-86.26
460021_3	<i>Ipomoea trifida</i>	CIP	2x	Venezuela		9	2	2	10.33	-67.66
460132_2	<i>Ipomoea trifida</i>	CIP	2x	Colombia		9			ND	ND
460039_1	<i>Ipomoea triloba</i>	CIP	2x	Colombia		3			3.41	-76.58
460039_2	<i>Ipomoea triloba</i>	CIP	2x	Colombia		3			3.41	-76.58
460044_1	<i>Ipomoea triloba</i>	CIP	2x	Colombia		3			4.96	-74.58
460044_2	<i>Ipomoea triloba</i>	CIP	2x	Colombia		3			4.96	-74.58
460116_1	<i>Ipomoea triloba</i>	CIP	2x	Colombia	40	3			10.5	-73.2
460116_2	<i>Ipomoea triloba</i>	CIP	2x	Colombia	39	3			10.5	-73.2
460658_1	<i>Ipomoea triloba</i>	CIP	2x	Colombia		18			4.11	-74.86
460658_2	<i>Ipomoea triloba</i>	CIP	2x	Colombia		18			4.11	-74.86
460101_1	<i>Ipomoea triloba</i>	CIP	2x	Colombia	4	22			11.36	-74.2
460101_2	<i>Ipomoea triloba</i>	CIP	2x	Colombia	5	22			11.36	-74.2
460526_2	<i>Ipomoea triloba</i>	CIP	2x	Cuba	36	3			21.75	-80
460090_1	<i>Ipomoea triloba</i>	CIP	2x	Dominican Republic		3			18.46	-69.96
460090_2	<i>Ipomoea triloba</i>	CIP	2x	Dominican Republic	34	3			18.46	-69.96
460100_1	<i>Ipomoea triloba</i>	CIP	2x	Ecuador		3			-2.16	-79.91
460518_2	<i>Ipomoea triloba</i>	CIP	2x	Ecuador	18	3			-1.66	-79.5
460782_1	<i>Ipomoea triloba</i>	CIP	2x	Ecuador	35	3			ND	ND
460782_2	<i>Ipomoea triloba</i>	CIP	2x	Ecuador	40	3			ND	ND
460518_1	<i>Ipomoea triloba</i>	CIP	2x	Ecuador	18	8			-1.66	-79.5
460554_1	<i>Ipomoea triloba</i>	CIP	2x	Mexico		18			18.66	-89.33
460554_2	<i>Ipomoea triloba</i>	CIP	2x	Mexico	3	18			18.66	-89.33
460558_1	<i>Ipomoea triloba</i>	CIP	2x	Peru	37	3			-12.58	-69.33
460558_2	<i>Ipomoea triloba</i>	CIP	2x	Peru	38	3			-12.58	-69.33
460012_1	<i>Ipomoea triloba</i>	CIP	2x	Venezuela		3			10.33	-67.66
460012_2	<i>Ipomoea triloba</i>	CIP	2x	Venezuela		3			10.33	-67.66
460052_1	<i>Ipomoea triloba</i>	CIP	2x	Venezuela		3			10.16	-66.66
460052_2	<i>Ipomoea triloba</i>	CIP	2x	Venezuela		3			10.16	-66.66

**Annexe 1: Données passeport des échantillons analysés dans les différents chapitres de cette étude.**

c) Echantillons d'herbiers

Sample number	Code	Country/archipelago	Date	Collector name	H	Cp lineage	Scoring quality	Missing data
BM_5	BM	Society Islands	1769	J. Banks and D. Solander	14	1	1	1
K_10	K	Easter Island	1911	F. Fuentes	14	1	2	1
K_11	K	Easter Island	1911	F. Fuentes	14	1	1	1
L_11	L	Austral Islands	1934	F. R. Fosberg	14	1	1	1
L_12	L	Austral Islands	1934	F. R. Fosberg	14	1	1	1
L_9	L	Pitcairn Island	1934	F. R. Fosberg and R. Clark	14	1	2	1
L_4	L	Caroline Islands	1946	E. Y. Hosaka	1_2	2	2	1
L_7	L	Fiji	1953	A. C. Smith	14	1	2	1
K_8	K	Gilbert and Ellice Islands	1968	F. Adair	14	1	2	1
L_15	L	Hawaii	1976	D. Herbst and S. Ishikawa	14	1	1	1
L_16	L	New Caledonia	1977	M. Mac Kee	14	1	2	1
R_39	BISH	Hawaii	1931	E.S. Handy	14	1	2	1
R_4	BM	Society Islands	1769	J. Banks and D. Solander	14	1	1	2
K_36	K	Madagascar	1882	G. W. Parker	14	1	2	2
K_37	K	Madagascar	1882	G. W. Parker	14	1	2	2
K_35	K	Madagascar	1896	ND	14	1	2	2
K_29	K	China	1898	A. Henry	1_2	2	2	2
K_13	K	Fiji	1949	L. V. Waibuta	1_2	2	2	2
K_1	K	New Zealand	1885	ND	14	1	2	3
BM_23	BM	Madagascar	1892	R. Baron	14	1	2	3
R_29	BISH	Hawaii	1931	E.S. Handy	14	1	1	3
L_8	L	Pitcairn Island	1934	F. R. Fosberg and R. Clark	14	1	2	3
L_5	L	Caroline Islands	1946	Y. C. Wong	1_2	2	2	3
L_6	L	Bismark Archipelago	1969	J. Rutherford	14	1	2	3
K_14	K	New Caledonia	1978	M. Mac Kee	14	1	2	3
BM_12	BM	Madeira	1837	J. Roch	1_2	2	2	4
BM_26	BM	Japan	1863	ND	1_2	2	2	4
K_23	K	Papua New Guinea	1889	William	14	1	2	4
R_31	BISH	Hawaii	1931	E.S. Handy	14	1	2	4
L_10	L	Gambier Islands	1934	F. R. Fosberg	14	1	2	4
BM_22	BM	Madeira	1768	J. Banks and D. Solander	6_9	2	2	5
BM_6	BM	New Zealand	1769	J. Banks and D. Solander	14	1	2	5
K_30	K	China	1889	A. Henry	1_2	2	2	5
K_34	K	Mauritius	1889	H.H. Jonhston	1_2	2	2	5
R_25	BISH	Hawaii	1937	F.E. Egler	1_2	2	2	5
R_2	BM	Carribean Islands	~1600	H. Sloane	1_2	2	2	5
K_9	K	Gambier Islands	1934	F. R. Fosberg	14	1	2	6
K_3	K	New Zealand	1935	L. A. Doore	14	1	2	6
K_20	K	Caroline Islands	1948	C. Y. C. Wong	1_2	2	2	6
K_12	K	Tuamotu islands	1990	J. F. G. Stokes	1_2	2	2	6
BM_10	BM	Philippines	1913	Merril	1_2	2		
R_32	BISH	Hawaii	1931	E.S. Handy	14	1		
L_2	L	Caroline Islands	1949	D. Anderson	14	1		
R_1	BM	Carribean Islands	~1600	H. Sloane	1_2	2		
R_3	BM	Carribean Islands	~1600	H. Sloane	1_2	2		
K_17	K	Hawaii	1862	ND	14	1		
R_33	BISH	Hawaii	1931	E.S. Handy	1_2	2		
R_34	BISH	Hawaii	1931	E.S. Handy	14	1		
R_35	BISH	Hawaii	1931	E.S. Handy	14	1		
R_45	BISH	Hawaii	1931	E.S. Handy	14	1		
K_2	K	New Zealand	1883	ND	14	1		
K_21	K	Papua New Guinea	1913	A. Pulle	1_2	2		
K_32	K	Philippines	1915	ND	14	1		
K_5	K	Samoa	1973	A. Whisler	1_2	2		
K_26	K	Timor	1883	A. B. Meyer	1_2	2		
K_6	K	Vanuatu	1971	H. S. McKee	6_9	2		
K_7	K	Vanuatu	1971	H. S. McKee	1_2	2		

## Annexe 2 : Index des noms de plantes cultivées cités dans le manuscrit.

Nom commun	Nom scientifique	Famille
Cordyline	<i>Cordyline fruticosa</i> L. (Chev.)	Agavaceae
Anone "cœur de bœuf"	<i>Anona reticulata</i> L.	Anonaceae
Alocase ou oreille d'éléphant	<i>Alocasia macrorrhiza</i> (L.) Schott.	Araceae
Macabo	<i>Xanthosoma sagittifolium</i> (L.) Schott	Araceae
Taro	<i>Colocasia esculenta</i>	Araceae
Taro géant	<i>Cyrtosperma chamissonis</i> (Schott.) Merrill	Araceae
Cocotier	<i>Cocos nucifera</i> L.	Arecaceae
Pejibayé	<i>Bactris gasipaes</i> Kunth	Arecaceae
Sagou	<i>Metroxylon sagu</i>	Arecaceae
Tournesol	<i>Helianthus annuus</i> L.	Asteraceae
Ananas	<i>Ananas comosus</i> (L.) Merr.	Bromeliaceae
Nangaille	<i>Canarium indicum</i> L.	Burseraceae
Patate douce	<i>Ipomoea batatas</i> (L.) Lam.	Convolvulaceae
Courge	<i>Cucurbita</i> sp.	Cucurbitaceae
Gourde / Calebasse	<i>Lagenaria siceraria</i> (Molina) Standl.	Cucurbitaceae
Grande igname	<i>Dioscorea alata</i> L.	Dioscoreaceae
Ignome africaine	<i>Dioscorea rotundata</i> Poir. et <i>Dioscorea cayenensis</i>	Dioscoreaceae
Ignome bulbifère	<i>Dioscorea bulbifera</i> L.	Dioscoreaceae
Ignome sauvage	<i>Dioscorea nummularia</i> Lam.	Dioscoreaceae
Manioc	<i>Manihot esculenta</i> Crantz	Euphorbiaceae
Arachide	<i>Arachis hypogaea</i> L.	Fabaceae
Châtaignier tahitien	<i>Inocarpus fagifer</i> (Parkinson) Fosberg	Fabaceae
Haricot commun	<i>Phaseolus vulgaris</i> L.	Fabaceae
Haricot de Lima	<i>Phaseolus lunatus</i> L.	Fabaceae
Choux des îles	<i>Abelmoschus manihot</i> (L.) Medik	Malvaceae
Coton	<i>Gossypium</i> sp.	Malvaceae
Arbre à pain	<i>Artocarpus altilis</i> (Parkinson) Fosberg	Moraceae
Mûrier à papier	<i>Broussonetia papyfera</i> (L.) Vent.	Moraceae
Bananier	<i>Musa</i> sp.	Musaceae
Oca	<i>Oxalis tuberosa</i> Molina	Oxalidaceae
Pandanus	<i>Pandanus</i> sp.	Pandanaceae
Kava	<i>Piper methysticum</i> Forst. f.	Piperaceae
Amidonnier	<i>Triticum turgidum</i> L.	Poaceae
Blé tendre	<i>Triticum aestivum</i> L.	Poaceae
Canne à sucre	<i>Saccharum officinarum</i> L.	Poaceae
Engrain	<i>Triticum monococcum</i> L.	Poaceae
Maïs	<i>Zea mays</i> subsp. <i>mays</i> L.	Poaceae
Millet perlé	<i>Pennisetum glaucum</i> (L.) R. BR .	Poaceae
Orge	<i>Hordeum vulgare</i> L.	Poaceae
Riz	<i>Oryza sativa</i> (subsp. <i>japonica</i> ou <i>indica</i> )	Poaceae
Poivron	<i>Capsicum</i> sp.	Solanaceae
Pomme de terre	<i>Solanum tuberosum</i> L.	Solanaceae
Tomate polynésienne	<i>Solanum repandum</i>	Solanaceae
Cacaoyer	<i>Theobroma cacao</i> L.	Sterculiaceae
Vigne	<i>Vitis vinifera</i> L.	Vitaceae

# Resumé

Au travers d'une approche de génétique des populations et de phylogéographie, basée sur la comparaison des patrons de diversité chloroplastiques et nucléaires, cette étude s'attache à retracer les processus qui ont façonné la diversité de la patate douce depuis sa domestication en Amérique tropicale jusqu'à son introduction et sa diffusion en Océanie.

Dans un premier temps, cette étude s'intéresse à l'histoire de la domestication et à l'identification de l'origine botanique et géographique de la patate douce dans son aire d'origine - l'Amérique tropicale. La patate douce est un taxon hexaploïde pour lequel les différents contributeurs sauvages n'ont pas encore été clairement identifiés. Deux hypothèses sont classiquement invoquées: 1) une origine autopolyploïde à partir d'un taxon sauvage diploïde *I. tridida* et 2) une origine allopolyploïde ayant impliqué l'hybridation de deux espèces distinctes *I. trifida* et *I. triloba*. Nos résultats génétiques viennent corroborer le scénario auto-polypléide. Néanmoins, contrairement à ce qui était avancé auparavant, *I. trifida* ne peut être considérée comme l'ancêtre sauvage de la patate douce. Des formes sauvages de *I. batatas* existent, populations à partir desquelles les formes cultivées ont été domestiquées. Par ailleurs, nous révélons l'existence de deux lignées chloroplastiques distinctes au sein des cultivars de patate douce, ce qui laisse penser que plusieurs parents sauvages, différenciés génétiquement mais probablement conspécifiques, sont impliqués dans la formation du génome de *I. batatas*. Deux scénarios (non exclusifs) peuvent alors être envisagés: i) *I. batatas* résulte de l'hybridation de plusieurs lignées distinctes (conspécifiques ou proches); ii) *I. batatas* est un complexe autopolyploïde avec une origine multiple. La caractérisation génétique des cultivars de patate douce met en lumière l'existence de deux groupes génétiques différenciés et géographiquement structurés: l'un correspond aux variétés d'Amérique centrale et caribéenne et l'autre aux variétés de la région du Pérou et de l'Equateur. Ce patron de diversité suggère fortement une domestication multi-locale - en Amérique centrale et en Amérique du Sud - et renforce l'hypothèse d'une origine autopolyploïde multiple dans ces deux régions.

Dans un deuxième temps, notre étude nous conduit en Océanie, une aire d'introduction de la patate douce. La distribution de la patate douce dans le Pacifique s'explique par une (des) introduction(s) pré-historique(s) en Polynésie en provenance d'Amérique du Sud (par les polynésiens eux-mêmes) et des introductions historiques dans le Pacifique Ouest, en provenance du Mexique et des Caraïbes. Il s'agit là d'une hypothèse élaborée par des linguistes, ethnobotanistes et archéologues, mais qui à ce jour manquait de preuves génétiques. En combinant un échantillonnage de variétés traditionnelles contemporaines et des spécimens d'herbiers datant du 18<sup>ième</sup> au début du 20<sup>ième</sup> siècle, nous avons pu retracer l'évolution temporelle et spatiale de la diversité dans le Pacifique. Nous montrons que les variétés de patate douce présentes jusqu'au début du 20<sup>ième</sup> siècle en Polynésie ont clairement une signature génétique sud-américaine, c'est-à-dire qu'elles dérivent directement des variétés de la région Pérou-Equateur. Ainsi nos données génétiques apportent une preuve supplémentaire à l'existence d'au moins une connexion préhistorique entre la Polynésie et l'Amérique du Sud. A l'Ouest du Pacifique, les cultivars de patate douce ont une origine principalement centraméricaine. Nous montrons également qu'il y a eu un remaniement de la base génétique au fil des nouvelles introductions, effaçant progressivement la signature des introductions d'origine. En revanche, les phénotypes reconnus par les cultivateurs et les noms associés - c'est-à-dire les déterminants « culturels » des variétés - ont probablement été maintenus. La patate douce est essentiellement propagée par voie clonale par les cultivateurs. Toutefois, elle a conservé une reproduction sexuée active. Nos données génétiques attestent que les nombreux cultivars qu'on trouve aujourd'hui en Océanie sont majoritairement issus d'événements de recombinaison distincts et de la sélection locale de plantes issues de graines. Dans certaines régions, ce processus de diversification a même conduit à l'émergence de véritables centres de diversité secondaire, comme dans les hautes terres de Nouvelle Guinée.

La patate douce est une plante alimentaire de première importance et la conservation des ressources génétiques une priorité. Les données obtenues dans notre étude permettent de mieux cerner la diversité « utile » de cette espèce. Aussi, l'histoire de sa diversification en Océanie est un modèle intéressant pour repenser la gestion et la conservation des plantes à racines et tubercules. Enfin, cette thèse confirme la pertinence d'étudier la diffusion des plantes cultivées pour retracer les migrations humaines et rappellent que l'élucidation de l'histoire croisée des hommes et des plantes est nécessairement un questionnement multidisciplinaire.

**Mots clés :** Patate douce, Océanie, domestication, polypléidie, propagation clonale, phylogéographie, agrobiodiversité.



# Summary

Following a population genetics and phylogeography approach, based on the comparison of chloroplastic and nuclear diversity patterns, this study aims at describing the processes which built sweet potato diversity from its domestication in tropical America to its introduction and diffusion into Oceania.

We first studied the history of sweet potato domestication and identified its botanic and geographic origin in the area from which it originates - tropical America. Sweet potato is a hexaploid taxa of which the wild parents still remain to be identified. Two hypothesis are classically referred to: 1) an autopolyploid origin deriving from a wild diploid *I. tridida* and 2) an allopolyploid origin implying the hybridization between *I. trifida* et *I. triloba*. Our genetic results corroborate the autopolyploid scenario. However, in contrast to what was previously anticipated, *I. trifida* cannot be considered the wild ancestor of sweet potato. Wild forms of *I. batatas* do exist, these are populations from which cultivated forms were domesticated. In addition, we highlighted the existence of two distinct chloroplastic lineages within sweet potato cultivars, suggesting that several wild parents, genetically differentiated but probably conspecifics are involved in the formation of the *I. batatas* genome. Two scenarios (non exclusive) are to be envisaged: i) *I. batatas* would result from the hybridization of several independent lines (conspecific or near); ii) *I. batatas* is an autopolyploid complex with multiple origin. The genetic characterization of sweet potato cultivars highlight the existence of two genetically differentiated and geographically structured groups: one includes central american and caribbean varieties while the other is made of varieties from Peru and Ecuador region. This diversity pattern is strongly suggestive of multilocal domestication events - in Central America and in South America - and strengthens the hypothesis of a multiple autopolyploid origin in these two regions.

Next, we investigated Oceania as area of introduction of sweet potato. The sweet potato distribution in the Pacific can be explained by pre-historic introductions in Polynesia originating from South America (brought by Polynesians), and historical introductions in West-Pacific originating from Mexico and the Caribbean islands. This is a hypothesis originally proposed by linguists, ethnobotanists and archeologists, but which was lacking until now of genetic proves. Combining the sampling of contemporary traditional varieties and herbarium specimens dating from the 18th to the early 20th century, we were able to refine the temporal and spatial evolution of sweet potato diversity in the Pacific. We demonstrate that sweet potato varieties present until the early 20th century in Polynesia clearly harbor a south-american genetic signature, indicating that they directly derive from varieties found in the Peru-Ecuador area. Thus our genetic data provide an additional prove to the existence of at least one prehistoric connection between Polynesia and South America. On the west side of Pacific, sweet potato cultivars mainly display a central-american origin. We also demonstrate that a reshuffling of the genetic base happened in line with the occurrence of new introductions, progressively erasing the signature of original introductions. In contrast, phenotypes and associated names known by farmers, i.e. cultural determinants of these varieties - were probably maintained over time. Sweet potato is essentially propagated clonally by farmers. However, it also maintained active sexual reproduction. Our genetic data demonstrate that the impressive numbers of cultivars found nowadays in Oceania mainly derive from independent recombination events and from the local selection of true-seed plants. In some regions, this diversification process even lead to the emergence of secondary diversity centers, as exemplified by New Guinea highlands.

Sweet potato is a food crop of high importance and the conservation of genetic resources remains a priority. Data obtained in our study enable to better scrutinize the "useful" diversity of this species. Also, the history of sweet potato in Oceania is an interesting model to re-evaluate the management and the conservation of root and tuber crops. Finally, this PhD work confirms the importance of studying the diffusion of cultivated plants to trace human migrations and highlight that elucidating the intertwined history of men and their plants is a mandatory multi-disciplinary questioning exercise.

**Key-words:** sweet potato, Oceania, domestication, polyploidy, clonal propagation, phylogeography, agrobiodiversity